DOCUMENT RESUME

ED 054 189 TE 499 819

AUTHOR	Thostenson, Marvin S.
TITLE	The Study and Evaluation of Certain Problems in Eartraining Related to Achievement in Sightsinging and Music Dictation.
INSTITUTION	Iowa Univ., Iowa City.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-6-8069
PUB DATE	Sep 66
CONTRACT	OEC-6-10-299
NOTE	62p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Auditory Training, *Comparative Analysis, *Music Reading, *Statistical Analysis, *Test Reliability
IDENTIFIERS	CSS76 Criterion Sightsinging Test, PRM78 Dictation Test

ABSTRACT

This investigation dealt with the development and evaluation of both a music dictation test (PRM78 Dictation Test) and a sightsinging test (CSS76 Criterion Sightsinging Test). It was hoped that the dictation test could eventually be developed to serve as an adequate replacement for the latter. Thirteen samples participated in this project--7 first-year undergraduate, 3 second and third year, and 3 graduate. Four research hypotheses were tested: (1) Achievement in the pitch, rhythm, and melodic sections of the PRM Dictation Test is substantially correlated with proficiency in the Criterion Sightsinging Test; (2) Achievement in the PRM Dictation Test is substantially correlated with proficiency in the Criterion Sightsinging Test; (3) Formal classwork in eartraining results in significant growth in music dictation achievement as measured by the PRM Dictation Test; and (4) The average achievement levels of first-year undergraduate students and other-than-first-year students in the two tests will favor the latter group somewhat, but are perhaps not of statistical significance. Results show that the PRM Dictation Test yielded total test reliabilities ranging generally from about .82 to .91, with a pitch section reliability of .86, rhythm .68, and melody .73. The CSS Criterion Sightsinging Test produced total rest reliabilities of about .95, with interval section of .84, pitch .93, rhythm .82, and melody .92. (DB)

ED054189

BR-6-8069

FINAL REPORT

Project No. 6-10-299

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE STUDY AND EVALUATION OF CERTAIN PROBLEMS IN EARTRAINING RELATED TO ACHIEVEMENT IN SIGHTSINGING AND MUSIC DICTATION

September 1966

TE 499 818

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

THE STUDY AND EVALUATION OF CERTAIN PROBLEMS IN EARTRAINING RELATED TO ACHIEVEMENT IN SIGHTSINGING AND MUSIC DICTATION

Contract No. 6-10-299

Marvin S. Thostenson

September 1966

The research reported herein was performed pursuant to a contract with the Office of Education, U. A. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of their project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

University of Iowa

Iowa City, Iowa

ACKNOWLEDGMENTS

This investigator is much indebted to many fellow teachers and a considerable number of students from several schools of music. But even more than just for the inconvenience that has been caused, we are indebted for the privilege it has been to investigate a very interesting area, under a sponsored project which has literally never had a dull moment.

We are pleased to express sincere gratitude to the following staff members from other colleges and universities:

Dr. Wayne Scott, University of Colorado, Boulder; Professor Gary Wittlich, University of Indiana, Bloomington; Dr. Maurice Monhardt, Luther College, Decorah, Iowa; Dr. Charles Spohn, Ohio State University, Columbus; Dr. Robert Jones, Southeast Missouri State College, Cape Girardeau; Dr. W. Kent Werner, Southern Illinois University, Carbondale; Dr. M. William Karlins, Western Illinois University, Macomb.

It has been a pleasure having the capable and reliable assistance of the following graduate students from the School of Music, University of Iowa:

Leland Bland, Christopher Bonds, Martin Farren, James Michel and Donald Pederson.

We acknowledge gratefully the professional counsel given by Professor John Hill, The School of Music, University of Iowa, concerning various phases of statistical analysis done in this project. The interest shown and the encouragement given by fellow staff members at the University of Iowa has not gone unnoticed. Further, the availability of the facilities of the Iowa Computer Center has been been of crucial importance.

Any acknowledgments are incomplete if they do not account for the many domestic adjustments that have been made both willingly and helpfully by my wife, Josephine.

Marvin S. Thostenson

CONTENTS

Title Page. i
Acknowledgements. ii
Contents. iii
List of Tables. iv

CHAPTERS

INTRODUCTION 1
Orientation and Problem 1
Background Research 1
Related Literature. 2
OBJECTIVES 5
METHOD 6
Testing-Music Dictation 6
Testing-Sightsinging. 7
Testing-Additional Samples. 8
Study of the Data 9
RESULTS. 12
The Individual Tests
The PRM78 Dictation Test--pre-test. . 12
The PRM72 Dictation Test--pre-test. . 14
The PRM78 Dictation Test--re-test . . 15
The CSS76 Criterion Sightsinging Test 19
The Pertinent Test Pairs. 22
Secondary Relationships 27
DISCUSSION 32
CONCLUSIONS. 38
SUMMARY. 42
REFERENCES 49

APPENDIXES

PRM78 Dictation Test A-1
Pitch Section A-2
Rhythm Section. A-5
Melodic Section A-7
CSS76 Criterion Sightsinging Test. B-1
Singing Intervals B-1
Singing Pitch Phrases B-2
Singing Rhythm Phrases. B-3
Singing Melodic Phrases B-4

LIST OF TABLES

Table		Page
I	-- Sample sizes for estimating growth in achievement in music dictation using PRM78	7
II	-- Sample sizes for estimating the relationship between music dictation, using PRM78 (pre-test and re-test), and sightsinging, using CSS76	7
III	-- Sample sizes for estimating the relationship between music dictation and sightsinging on both the first-year and other-than-first-year levels	8
The PRM78 Music Dictation Test		
-----Pre-test		
IV	-- Central tendency and dispersion for the total PRM78 Dictation pre-test, 6 samples	13
V	-- Four estimates of reliability--test sections and the total test (pitch--30 items; rhythm--24 items; melody--24 items; total test--78 items)	13
VI	-- Within-test inter-correlations for six first-year samples	14
VII	-- Comparison of four estimates of reliability--test sections and total test, PRM78 and PRM72	15
-----Re-test		
VIII	-- Homogeneity of variance analysis	16
IX	-- Central tendency and dispersion on the total test, 13 samples	16
X	-- Four estimates of reliability--test sections and total test, 13 samples	16
XI	-- The within-test inter-correlations for 13 samples as listed	19

Table | Page

The CSS76 Criterion Sightsinging Test

XII -- Homogeneity of variance 19

XIII -- Central tendency and dispersion of the total test, CSS76, 7 samples 20

XIV -- Four estimates of reliability--test sections and the total test (singing intervals--12 items; singing pitch phrases--24 items; singing rhythm phrases--20 items; singing melodic phrases--20 items; total test--76 items) 20

XV -- Within-test correlations for the samples listed separately, and in certain combinations 22

The Pertinent Test Pairs

-----PRM78 Dictation Test--pre-test and re-test

XVI -- The significance of the difference between correlated means (reliabilities; mean gain; "t" ratio; and interpretation) 22

-----PRM78 Dictation re-test and the CSS76 Criterion Sightsinging Test

XVII -- The correlation between music dictation and sightsinging 24

XVIII -- Inter-correlations between PRM78 Dictation re-test and CSS76 Criterion Sightsinging Test--four first-year samples, N=147 25

-----PRM78 Dictation pre-test and CSS76 Criterion Sightsinging Test

XIX -- Correlation of music dictation (early Fall 1965) with sightsinging (Spring 1966)--four first-year samples, N=147 26

Table | Page

-----PRM78 Dictation re-test and CSS76 Criterion Sightsinging Test

XX -- Inter-correlations between PRM78 Dictation re-test and CSS76 Criterion Sightsinging Test--four first-year samples, N=147 27

Secondary Relationships

XXI -- The relationships between type of degree sought and achievement in music dictation and sightsinging (*=.05 and **=.01 levels for "F" and "t") 28

XXII -- The relationships between the area of instrumental major and achievement in music dictation and sightsinging (*=.05 and **=.01 levels for "F" and "t" ratios) 29

XXIII -- The relationships between training on major instrument and achievement in music dictation and sightsinging (*=.05 and **=.01 levels for "F" and "t" ratios) 31

INTRODUCTION

ORIENTATION AND PROBLEM

Many Schools of Music are concerned about not only the quality of students with whom they are working, but also concerning their college level courses for the training of these musicians. Even though this is perhaps particularly true in the area of eartraining, there is also considerable evidence of interest in all the areas of information and skill pertinent to the proper development of the music major on all levels of training. The principal objectives of this investigation were to estimate the growth in dictation proficiency achieved in the course of the first year of eartraining, to estimate the relationship between proficiency in music dictation and in sightsinging at both the undergraduate and the graduate levels, and further, to note the inter-relationships between these proficiencies under the varying conditions associated with the type of music degree being sought, the area of the major instrument, and the amount of training on the major instrument, measured in terms of the years of private study.

BACKGROUND RESEARCH

On the basis of testing conducted between January 1964 and September 1965, this investigator constructed a thirty-minute sightsinging test consisting of four sections, namely, 1) singing the 12 simple intervals, 2) singing 24 four-note pitch phrases without rhythm obligation, 3) singing 20 two-bar rhythm phrases without pitch obligation, and 4) singing 20 two-bar melodic phrases with both pitch and rhythm obligations. During the first three months of this project some selective shortening of the second and some revision of the third and fourth sections was accomplished in an effort to improve, in part, the objectivity of the test. These steps were further intended to make test evaluation among the four administrators as uniform as was possible. Since the sightsinging test was comprehensive, individually administered, and was designed to measure a skill whose possession is quite generally accepted as

one indication of functional musicianship, this 76 item test was designated as the CSS76 Criterion Sightsinging Test.

Because of the importance of estimating proficiency in sightsinging, while concurrently being aware of the considerable inconvenience and even impracticality of doing such thorough testing continuously on an individual basis, it was hoped that an approximately equivalent, objective, group music dictation test could be constructed. The instrument developed during the same prior period of study and testing was the PRM78, Pitch, Rhythm, Melodic Dictation Test. As used in this research project, the dictation test consisted of 78 four-foil multiple choice items. These included 1) 30 pitch items constructed as short melodic phrases whose rhythm factor was unchanging, 2) 24 rhythm items constructed as short melodic phrases whose pitch factor was unchanging, and 3) 24 melodic items constructed as short phrases whose pitch and rhythm content each differed from foil to foil. The student's task in listening to the tape recording of the test was to select for each item played the foil whose notation correctly represented the sound that was heard.

In the Spring of 1965 the immediate predecessors of the dictation and sightsinging tests used in the current research project were given to 54 first-year students in the School of Music of the University of Iowa. These tests yielded respective whole test reliabilities of .84 and .94, together with an encouraging inter-correlation coefficient of..86.

RELATED LITERATURE

Ever since the eleventh century when Guido devised the first system of solmization syllables there has been interest in sightsinging. Randall Thompson states that "sightsinging is meant to enable a student to hear a written melody without recourse to an instrument, (and) eartraining is meant to teach him to write down what he hears. Each is a reversal of the processes of the other; both are considered indispensible to the study of theory." (18)

Perhaps the larger number of studies have concerned the prediction of success in music through attempting to measure musical aptitude. These range from the Seashore view with stress on measuring acoustical-sensory capacities, (16) to the Gestalt approach of attempting to assess total musicality. (14) Although the Seashore tests have enjoyed wide use for many years, the Gestalt view of Revész concerning the total musical person has been favored by other writers and researchers such as Drake, (3), Mursell, (10), Wing, (19), Lundin, (6), and Gordon. (5)

Even though sightsinging is a skill of some complication it must be considered to be only a part of the total concept of musicality in the opinion of Ottman. (12) Perhaps the key word is "skill". Studies in musical aptitude tend to avoid tests involving notation; therefore sightsinging and music dictation--as skills involving the reading and interpreting of musical symbols--are, rather, tests in musical achievement. In this study the relationships that undoubtedly do exist between musical aptitude and achievement in sightsinging and music dictation are left quiescent and are not matters for consideration or evaluation.

Important basic research in music dictation was done by Ortmann some 30 years ago, (11) in which he concluded that achievement in music dictation was inhibited by student problems in music notation. Madison contributed to the knowledge concerning interval hearing and found this skill significantly related to musicality. (7), Forty years ago Moser investigated the relationship between sightsinging and several other musical achievements, deriving the highest correlation, .62, between sightsinging and the dictation of tonal patterns. (9) Dean in using both aptitude and achievement testing to predict success in sightsinging, found correlations of .61 and .63 between sightsinging and the Seashore pitch and tonal memory subtests, and concluded that other separate factors were of little value in the area. (2) Earlier Salibury and Smith had teamed these tests with melodic dictation to achieve their best correlation with sightsinging and to derive a predictive regression equation. (15) In a more recent

study of considerable importance, Ottman assessed the correlation of a number of skills and attributes with sightsinging ability. Among these he found a correlation between sightsinging and a set of interval tests to be .68, and a correlation of .73 with his "music literacy" test (in which error detection is used). He suggested that the latter area and the investigation of rhythm problems of sightsinging (both of which are important aspects of the present project) were worthy of considerable study. (12) Aliferis has made available multiple choice dictation tests with which to measure aural achievement in melodic, harmonic, and rhythm concepts at the entrance and midpoint levels of undergraduate study, without, however, defining any relation to proficiency in sightsinging. (1) The entrance test is valuable as an initial eartraining sectioning device, but uses rather extensive range and complicated notation to achieve satisfactory discrimination and thereby possibly introduces some extraneous factors for this early stage. William Poland enlarged upon Ortmann's research by demonstrating that knowledge of fundamental concepts is essential for effective work in undergraduate written and aural music theory. (13) A recent investigation by Marquis suggested that isolated intervals functioned differently from intervals influenced by context in melodic sightsinging, but also reaffirmed that the ability to sing isolated intervals was substantially correlated with the ability to sightsing. (8) Among the more important research being conducted presently is that of Charles Spohn of Ohio State University, Paul Harder and Merrill Sherburn of Michigan State University, and James Carlson and Walter Ihrke of the University of Connecticut. In a recent research report Spohn found significant differences in improvement for interval and rhythm eartraining stemming from differences in pedagogical approach, comparing four taped programmed laboratory procedures. (17) In both concepts--intervals and rhythm--the best method produced excellent results. The work at Michigan State University lies principally in the area of programmed and taped instruction. That at the University of Connecticut covers aspects of the same area under the general title of "Automated Music Instruction."

This brief review suggests something of the breadth of the problem, several areas of which are given some attention in this report.

OBJECTIVES

The immediate purpose of this project was to investigate problems concerning several skills related to music dictation and sightsinging. The investigation dealt with the study and evaluation of music dictation and sightsinging achievement as they exist and develop under the essentially non-controlled conditions of classroom teaching found in most music schools. Each of the immediate objectives is also stated as an hypothesis. The study was designed--

1. to estimate the correlation of the separate sections of the Pitch, Rhythm, and Melodic Dictation Test with proficiency in the Criterion Sightsinging Test.

Research Hypothesis #1: Achievement in the pitch, rhythm, and melodic sections of the PRM Dictation Test is substantially correlated with proficiency in the Criterion Sightsinging Test.

2. to estimate the correlation of the PRM Dictation Test with proficiency in the Criterion Sightsinging Test.

Research Hypothesis #2: Achievement in the PRM Dictation Test is substantially correlated with proficiency in the Criterion Sightsinging Test.

3. to assess the growth in music dictation achievement resulting from the formal study of eartraining by first-year students.

Research Hypothesis #3: Formal classwork in eartraining results in significant growth in music dictation achievement as measured by the PRM Dictation Test.

4. to compare the average achievement levels in music dictation and in sightsinging attained by first-year and other-than-first-year students.

Research Hypothesis #4: The average achievement levels of first-year undergraduate students and other-than-first-year students in the PRM Dictation Test and

the Criterion Sightsinging Test will favor the latter group somewhat, but perhaps by an amount too small to achieve statistical significance.

The original statement used in the proposal was "to compare the average achievement levels in music dictation and sightsinging of the other-than-first-year undergraduate and graduate students of the University of Iowa School of Music." It was found impossible to accomplish this testing in its entirety due to a necessary two-month shortening of the project's duration. However, the testing of two graduate samples from the University of Iowa totalling 133 students, together with a group of 15 sophomores taking remedial eartraining, furnished a 148 student sample which, together with the 147 first-year students from four samples, constituted in our opinion an adequate test of the intent of this hypothesis.

A secondary purpose of this research project was to study other relationships which developed logically from the data that were collected. The test data were studied under three conditions. These were the interrelationships between proficiency in music dictation and sightsinging and 1) the type of degree being sought, 2) the category of the major instrument, and 3) the amount of private instruction on the major instrument.

METHOD

TESTING -- MUSIC DICTATION

In the early Fall of 1965 the first-year music students of five college and university schools of music were given the PRM78 Dictation Test. At the end of the first semester a sixth similar sample was added. In May 1966 the end-of-first-year students of the same six schools were retested. The six samples available for estimating the growth in music dictation proficiency achieved during the first year of eartraining classwork as measured by the PRM78 Dictation Test are listed. This group of samples was used to test the truth of Hypothesis #3 relating to growth in dictation achievement, stated above under "Objectives."

TABLE I -- SAMPLE SIZES FOR ESTIMATING GROWTH IN ACHIEVEMENT IN MUSIC DICTATION USING PRM78

Sample	Pre-test	Re-test	Usable Sample*
UG#1	90	65	62
UG#2	43	23	17
UG#3	75	50	47
UG#4	76	40	40
UG#5	69	30	30
UG#6	62	60	51
TOTAL	415	268	247

* Students completing both the pre-test and the re-test, together with other needed information. Several factors, but principally attrition contributed to the reduction in sample sizes available.

TESTING -- SIGHTSINGING

In May 1966 the Criterion Sightsinging Test, CSS76, was given to the first-year students from the four schools that participated in this phase of the project. This individual test was given to the students available at the school site by the graduate assistants working with this investigator. This group of first-year samples provided data for testing certain aspects of Hypotheses #1 and #2 relating to the correlation between music dictation and sightsinging, and Hypothesis #4 relating to differences in achievement in music dictation and sightsinging between different groupings of students.

TABLE II -- SAMPLE SIZES FOR ESTIMATING THE RELATIONSHIP BETWEEN MUSIC DICTATION, USING PRM78 (pre-test and re-test), AND SIGHTSINGING, USING CSS76

Sample	Pre-test PRM78 (pre)	Re-test PPM78 (re)	S's'g Test CSS76	Usable Sample
UG#1	90	65	62	62
UG#2	43	23	17	15
UG#3	75	50	47	41
UG#4	76	40	39	29
TOTAL	284	178	165	147

During the period from March through June 1966 97 graduate students, 36 special graduate students enrolled in an Arts and Humanities Summer Workshop in the Department of Music Education, and 15 sophomores completing the fourth semester of eartraining, all in the School of Music of the University of Iowa, were given both the PRM78 Dictation Test and the CSS76 Criterion Sightsinging Test. Therefore, for the comparison of attainment on the dictation test with that measured at about the same time on the sightsinging test, seven population samples were available. These seven samples are listed in Table III. The evaluation of the data from these samples provided for further testing of the truth of Hypothesis #1, #2, and #4.

TABLE III -- SAMPLE SIZES FOR ESTIMATING THE RELATIONSHIP BETWEEN MUSIC DICTATION AND SIGHTSINGING ON BOTH THE FIRST-YEAR AND OTHER-THAN-FIRST-YEAR LEVELS.

Sample	PRM78	CSS76	Usable Sample	Remarks
UG#1	65	62	62	
UG#2	23	17	15	
UG#3	50	47	41	
UG#4	40	39	31	
UG#8	15	15	15	U of Iowa sophomores in 4th semester eartraining
G#1	97	97	97	U of Iowa graduate students, about 50% of graduate enrollment, 1965-66
G#2	36	36	36	U of Iowa, Arts and Humanities Music Education Workshop
TOTAL	326	313	297	

TESTING -- ADDITIONAL SAMPLES

During May 1966 a seventh school added one more substantial first-year sample (UG#7, N=94), and another provided a small sample of graduate students (G#3, N=19).

The school providing first-year sample, UG#3, also made available a sophomore sample (UG#10, N=28) and an additional group of 28 sophomores and juniors was procured from the University of Iowa. These added samples made available data concerning only the PRM78 Dictation Test, but this assisted materially in testing the intent of Hypothesis #4, which premised that differences in achievement levels in music dictation and in sight-singing are not as different between first-year students and other-than-first-year students as one might be inclined to suppose.

STUDY OF DATA

Two test instruments were utilized in this project, namely, the PRM78 Dictation Test and the CSS76 Criterion Sightsinging Test.

The PRM78 Dictation Test was studied under two temporal conditions relating to first-year music theory students, namely, 1) as a pre-test given in the Fall of 1965 (five samples then and one at the end of the first semester--total of six samples), and 2) as a re-test given in the Spring of 1966 (the above six samples, one more first-year sample, three other non-first-year undergraduate samples, and three graduate samples--total of 13 samples). Further PRM78 was studied under two analytical conditions, namely, 1) as the total test of 78 items of pitch, rhythm, and melody, and 2) as a 72 item test by eliminating from statistical study six items considered to be weak.

The CSS76 Criterion Sightsinging Test was given under only one temporal condition, that is, near the close of the school year, and was not studied in an abbreviated form.

The data for the PRM78 Dictation Test, that for the abbreviated 72 item dictation test--each of these as a pre-test and as a re-test, and, the CSS76 Criterion Sightsinging Test were analyzed in several ways.

1. Each was considered as a separate testing instrument and the following statistics were derived for each sample studied:

a. sample size
b. mean score
c. SE of the mean
d. SD of the population concerned, estimated from the sample
e. the test reliability was estimated using four different statistical approaches--
1) the Kuder-Richardson Approximation Formula
2) the Kuder-Richardson Internal Consistency Formula
3) applying the Spearman-Brown Prophecy Formula to the Pearson product-moment correlation of the two test halves based on an ordinary odd-even split-half ordering
4) applying the Spearman-Brown Prophecy Formula to the Pearson product-moment correlation of the two test halves based on a selective ordering of the test items. This order was based upon their value as ascertained by item analysis study of the data from the 97-student graduate sample from the University of Iowa. This particular ordering was used in determining this reliability coefficient from each sample studied.
f. the inter-correlations between the test sections and that between each section and the total test.

2. The tests of a given sample were studied for specific purposes pertinent to the investigation.
a. the PRM78 Dictation pre-test in relation to the PRM78 Dictation re-test for examining Hypothesis #3 regarding the change in music dictation achievement over a school year of eartraining study.
b. the PRM78 Dictation re-test given in the Spring of 1966, and the CSS76 Criterion Sightsinging Test given at about the same time for examining the truth of Hypotheses #1 and #2, concerning inter-correlations between music dictation and sightsinging, and Hypothesis #4 relating to differences in achievement between first-year and other-than-first-year levels.
c. the PRM78 pre-test given in the Fall of 1965 (UG samples #1-#4) and the CSS76 Criterion Sightsinging Test given to the same students in the Spring of 1966. Although this validity correlation was not mentioned in the original proposal, its computation was a natural result of the testing program followed and the results

constitute an important addition to this completed project's worth.

3. After computing these statistics for each sample it was necessary to establish whether the samples could be combined for further study. To this end the Bartlett Test of the Homogeneity of Variance was applied to the following sample groups: (4)

a. the PRM78 Dictation Test
   1) four first-year samples on the pre-test
   2) four first-year samples on the re-test
   3) six first-year samples on the pre-test
   4) six first-year samples on the re-test
   5) cumulatively, seven first-year samples on the re-test
   6) three graduate samples on the re-test
   7) cumulatively, 13 undergraduate and graduate samples on re-test

b. the CSS76 Criterion Sightsinging Test
   1) four first-year samples
   2) one sophomore sample and two graduate samples
   3) cumulatively, seven undergraduate and graduate samples.

4. Assuming that the samples could be proved homogeneous in variance by Bartlett's Test it would next be in order to study the three basis pairs of tests mentioned above.

a. the PRM78 Dictation pre-test and re-test
b. the PRM78 re-test and the CSS76 Criterion Sightsinging Test, both given in the Spring of 1966
c. the PRM78 pre-test given near the beginning of formal eartraining and the CSS76 Criterion Sightsinging Test given near the end of the first year of formal eartraining

For each of these three pairs of tests the following statistics were compiled from the data available:

1) the sample size, or composite sample size
2) the homogeneity between samples
3) the mean score
4) the SE of each mean score
5) the estimate of the population SD based on the given sample

6) the four estimates of reliability
7) the inter-correlations within each test
8) the significance of the difference between means for each section (pitch, rhythm, and melody) and the whole test.

Relative to all three pairs of tests in each combination studied the following information was made available:

1) the inter-correlations within and between the several test sections and between the total test scores of the two tests concerned
2) the measure of the between-section and the between-test differences in variance as expressed through the F ratio
3) the interpretation and location of significant variance accumulations through computing the pertinent "t" ratios.

5. The relationship between achievement in music dictation and sightsinging and the three following conditions, namely, 1) the type of degree sought, 2) the area of the major instrument, and 3) the extent of formal study on the major instrument.

RESULTS

THE INDIVIDUAL TESTS

The PRM78 Dictation Test -- pre-test

1. The Homogeneity of Variance

Bartlett's Test of the homogeneity of variance applied to the four first-year samples gave an uncorrected* chi-square of 4.94, whereas a chi-square of 7.82 was necessary for difference at the five-percent level with three degrees of freedom. Six first-year samples made available a chi-square of 6.29 with one at 11.07 needed for difference at the five-percent level with five degrees of freedom. Therefore, the six samples were treated as being homogeneous, that is, deriving from the same parent population of students.

* The correction slightly reduces the uncorrected value.

2. TABLE IV -- CENTRAL TENDENCY AND DISPERSION FOR THE TOTAL PRM78 DICTATION PRE-TEST, 6 SAMPLES

Sample	N	Mean	SE	SD(pop)
UG#1	62	45.76	1.70	13.36
UG#2	17	36.53	2.65	10.92
UG#3	47	45.00	1.66	11.08
UG#4	40	39.10	1.40	8.87
UG#5	30	37.63	1.68	9.22
UG#6	51	44.02	1.69	12.07
TOTAL	247	42.55	.75	11.78

3. TABLE V -- FOUR ESTIMATES OF RELIABILITY -- TEST SECTIONS AND TOTAL TEST (Pitch--30 items; Rhythm--24 items; Melody--24 items; total test--78 items)

Sample-N	TEST	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#1	P	.8634	.8736	.8321	.8876	.27
N=90	R	.6430	.6873	.6722	.6382	
	M	.7375	.7525	.7002	.7553	
	PRM	.8945	.9041	.8903	.9127	
UG#2	P	.7141	.7349	.7335	.7798	.39
N=43	R	.6301	.6658	.7603	.6404	
	M	.4987	.5292	.6676	.6002	
	PRM	.8063	.8212	.8242	.8753	
UG#3	P	.8481	.8616	.8682	.9009	.30
N=75	R	.5423	.5924	.6936	.6028	
	M	.6887	.7047	.7776	.6849	
	PRM	.8584	.8721	.8953	.8999	
UG#4	P	.7460	.7706	.7690	.8271	.30
N=76	R	.5864	.6271	.7217	.6936	
	M	.5958	.6241	.5347	.5918	
	PRM	.8061	.8236	.7797	.8481	
UG#5	P	.6541	.6807	.7071	.7844	.30
N=69	R	.6089	.6592	.6134	.6669	
	M	.5435	.5781	.6579	.6698	
	PRM	.7737	.7957	.8507	.7826	

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#6	P	.8423	.8583	.9081	.8834	.33
N=60	R	.4752	.5416	.5835	.4959	
	M	.6963	.7189	.7569	.8184	
	PRM	.8551	.8693	.8863	.8837	
UG#1-#4	P	.8338	.8459	.8303	.8766	.16
N=284	R	.6777	.7087	.7481	.7120	
	M	.7104	.7269	.7033	.7198	
	PRM	.8802	.8896	.8836	.9078	
UG#1-#6	P	.8254	.8377	.8379	.8725	.13
N=415	R	.6500	.6850	.7128	.6842	
	M	.6962	.7140	.7086	.7337	
	PRM	.8707	.8809	.8845	.8942	

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

4. TABLE VI -- WITH-IN TEST INTER-CORRELATION FOR SIX FIRST-YEAR SAMPLES

Sample	N	P/R	P/M	P/PRM	R/M	R/PRM	M/PRM	"r".01*
UG#1	90	.45	.73	.92	.50	.71	.88	.27
UG#2	43	.24	.71	.92	.16	.51	.83	.39
UG#3	75	.06	.65	.89	.32	.42	.87	.30
UG#4	76	.11	.47	.70	.45	.57	.87	.30
UG#5	69	.41	.43	.82	.38	.75	.77	.30
UG#6	60	.30	.68	.89	.43	.62	.87	.33
UG#1-#4	284	.27	.70	.89	.45	.61	.89	.16
UG#1-#6	415	.26	.66	.88	.45	.62	.87	.13

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

The PRM72 Dictation Test -- pre-test

The PRM78 Dictation Test was reduced to PRM72 by eliminating from the computations six items of data that were thought to be too unreliable, as mentioned previously. Since the statistics computed did not show any effective increase in reliability, further use of this

information was abandoned. However, this conclusion was reached after checking the actual results obtained from a PRM72 analysis and a PRM78 analysis for each sample concerned. The following table showing one set of comparisons is the only one given in this report.

TABLE VII -- COMPARISON OF FOUR ESTIMATES OF RELIABILITY -- TEST SECTIONS AND TOTAL TEST, PRM78 AND PRM72

Sample	N	Test	KRAP	KRIC	ORDSH	SELSH	Remarks
UG#1-6	415	P30	.8254	.8377	.8379	.8725	30 pitch items in both PRM78 and PRM72
		R24	.6500	.6850	.7128	.6842	24 items in PRM78
		R20	.6141	.6525	.6805	.6153	20 items in PRM72
		M24	.6962	.7140	.7086	.7337	24 items in PRM78
		M22	.6699	.6878	.7150	.7213	22 items in PRM72
		PRM78	.8707	.8808	.8845	.8942	78 item test
		PRM72	.8653	.8755	.8779	.8896	72 item test

The PRM78 Dictation Test -- re-test

Seven first-year samples were included in this phase of study. The word "re-test" is defined as meaning the administration of the PRM78 Dictation Test after a year of training for first-year classes, independent of whether or not the pre-test was given. However, since any group beyond the first year is relatively, a "trained group", all samples participating in the project at this time are included in this portion of the report.

1. The Homogeneity of Variance

Bartlett's Test of the homogeneity of variance produced the following results, and these indicate that each combination of samples is representative of the same general student population of music students. This statement applies to both the undergraduate and graduate students.

TABLE VIII -- HOMOGENEITY OF VARIANCE ANALYSIS

Group	No. of Samples	N	Degrees of Freedom	Uncorrected Chi-square	Chi-square Sig. at .05
1st-year students	7 from 7 schools	371	364	9.13	12.59 at 6df
Graduate students	3 from 2 schools	152	149	.760	5.99 at 2df
Soph. & Juniors	3 from 2 schools	71	68	4.37	5.99 at 2df
All samples	13 from 8 schools	594	581	18.25	21.03 at 12df

2. TABLE IX -- CENTRAL TENDENCY AND DISPERSION ON THE TOTAL TEST, PRM78 RE-TEST, 13 SAMPLES

Sample	N	Mean	SE	SD(pop)
UG#1	65	52.22	1.46	11.75
UG#2	23	42.35	2.54	12.20
UG#3	50	50.86	1.68	11.87
UG#4	49	45.76	1.79	12.53
UG#5	30	41.77	1.71	9.36
UG#6	60	46.55	1.71	12.19
UG#1-6	247	48.13	.78	12.31
UG#7	94	49.30	1.01	9.78
UG#1-7	341	48.04	.60	11.61
UG#8	15	49.47	3.05	11.83
UG#9	28	50.86	1.63	8.61
UG#10	28	53.18	2.26	11.94
G#1	97	53.22	1.72	10.92
G#2	36	55.44	1.96	11.77
G#3	19	54.26	2.44	10.65

3. TABLE X -- FOUR ESTIMATES OF RELIABILITY -- TEST SECTIONS AND TOTAL TEST, PRM78 RE-TEST, 13 SAMPLES

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#1	P	.8586	.8716	.8997	.8621	.33
N=65	R	.6810	.7162	.7054	.8203	
	M	.6288	.6535	.7311	.7361	
	PRM	.8843	.8956	.8990	.9138	

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#2	P	.8440	.8590	.8893	.9242	.50
N=23	R	.5781	.6452	.7030	.7329	
	M	.6887	.7109	.8061	.7537	
	PRM	.8752	.8887	.9312	.9308	
UG#3	P	.8930	.9000	.9290	.8459	.36
N=41	R	.5387	.6020	.5709	.5700	
	M	.7093	.7275	.6536	.6867	
	PRM	.8832	.8933	.8778	.8501	
UG#4	P	.8608	.8743	.9268	.9121	.41
N=39	R	.5968	.6426	.4731	.5760	
	M	.6894	.7110	.6504	.7095	
	PRM	.8884	.8991	.8801	.8901	
UG#1-4	P	.8715	.8805	.9141	.8784	.20
N=168	R	.6560	.6914	.6401	.7165	
	M	.7133	.7281	.7250	.7432	
	PRM	.8936	.9018	.8949	.9008	
UG#5	P	.5402	.5811	.6780	.6766	.44
N=30	R	.5921	.6638	.6926	.4674	
	M	.5675	.6075	.7029	.6252	
	PRM	.7810	.8127	.8763	.8299	
UG#6	P	.8447	.8605	.9117	.8227	.33
N=60	R	.6546	.6792	.7121	.7463	
	M	.6412	.6658	.6594	.6290	
	PRM	.8671	.8782	.9158	.8608	
UG#1-6	P	.8594	.8693	.9082	.8642	.16
N=258	R	.6526	.6849	.6615	.6984	
	M	.6961	.7129	.7176	.7166	
	PRM	.8853	.8941	.9009	.8903	
UG#7	P	.8166	.8316	.8565	.8164	.26
N=94	R	.3540	.4232	.4365	.5074	
	M	.5002	.5352	.6314	.4627	
	PRM	.8189	.8361	.8740	.8319	

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#1-7	P	.8503	.8606	.8972	.8523	.14
N=352	R	.6095	.6458	.6277	.6697	
	M	.6644	.6837	.7043	.6748	
	PRM	.8739	.8838	.8961	.8802	
UG#8	P	.8059	.8339	.8609	.8750	.58
N=15	R	.3080	.4336	.5854	.6685	
	M	.3213	.3827	.2625	.6738	
	PRM	.6976	.7411	.7122	.7767	
UG#9	P	.8143	.8326	.8220	.8719	.44
N=28	R	.6372	.6966	.7306	.7638	
	M	.4390	.4835	.4066	.5050	
	PRM	.7624	.7879	.7367	.7819	
UG#10	P	.8359	.8467	.8395	.8450	.44
N=28	R	.5352	.6181	.4785	.5561	
	M	.8108	.8306	.7281	.8592	
	PRM	.8884	.9005	.9148	.9430	
G#1	P	.8449	.8553	.8721	.8448	.26
N=97	R	.5941	.6403	.5700	.6412	
	M	.7251	.7443	.7534	.7835	
	PRM	.8969	.9139	.9138	.9020	
G#2	P	.8821	.8922	.8919	.8931	.42
N=36	R	.7063	.7390	.7940	.6813	
	M	.7225	.7485	.7793	.8292	
	PRM	.8924	.9021	.9208	.9214	
G#3	P	.8673	.8857	.8527	.9031	.53
N=19	R	.4778	.6172	.4450	.8005	
	M	.6531	.6974	.6490	.6510	
	PRM	.8573	.8814	.8331	.8493	

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

TABLE XI --

4. THE WITHIN-TEST INTER-CORRELATIONS FOR 13 SAMPLES AS LISTED

Sample-N	P/R	P/M	P/PRM	R/M	R/PRM	M/PRM	"r".01*
UG#1-65	.51	.70	.93	.38	.72	.82	.32
UG#2-23	.51	.65	.92	.42	.72	.83	.50
UG#3-50	.19	.73	.92	.26	.48	.87	.36
UG#4-49	.39	.78	.91	.62	.71	.93	.36
UG#1-4	.27	.70	.89	.45	.60	.89	.19
N=187							
UG#5-30	.44	.56	.81	.62	.81	.87	.44
UG#6-60	.32	.65	.88	.44	.66	.84	.33
UG#1-6	.39	.73	.91	.46	.68	.87	.16
N=277							
UG#7-94	.33	.69	.93	.25	.57	.82	.26
UG#1-7	.38	.70	.91	.44	.67	.86	.14
N=371							
UG#8-15	.04	.20	.84	.15	.43	.87	.58
UG#9-28	.00	.49	.86	.07	.41	.66	.44
UG#10-28	.29	.68	.86	.64	.69	.93	.44
G#1-97	.26	.73	.90	.27	.57	.86	.26
G#2-36	.31	.56	.86	.54	.70	.84	.42
G#3-19	.11	.73	.89	.35	.49	.90	.53

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

The CSS76 Criterion Sightsinging Test

1. The Homogeneity of Variance

Again the use of Bartlett's Test revealed that the variance among the seven samples participating in this phase of the project differed only within the limitations of chance. The table gives the statistics derived from the several groupings of samples.

TABLE XII -- THE HOMOGENEITY OF VARIANCE

Group	No. of Samples	N	Degrees of Freedom	Uncorrected Chi-square	Chi-square Sig. at .05
1st-year students	4 from 4 schools	147	143	.889	7.815 with 3df

Group	No. of Samples	N	Degrees of Freedom	Uncorrected Chi-square	Chi-square Sig. at.05
Other-than 1st-year	3 from 2 schools	148	145	2.685	5.991 with 2df
Total of samples	7 from 4 schools	295	288	3.603	12.592 with 6df

TABLE XIII -- CENTRAL TENDENCY AND DISPERSION FOR THE TOTAL TEST, CSS76, 7 SAMPLES

Sample	N	Mean	SE	SD(pop)
UG#1	62	36.85	2.14	16.84
UG#2	15	22.27	4.28	16.59
UG#3	41	35.05	2.64	16.89
UG#4	29	26.87	2.70	14.52
UG#1-4	147	32.90	1.41	17.04
UG#8	15	33.87	3.05	11.83
G#1	97	42.63	1.72	16.91
G#2	36	45.14	2.62	15.74

TABLE XIV -- FOUR ESTIMATES OF RELIABILITY -- TEST SECTIONS AND THE TOTAL TEST (Singing Intervals--12 items; Singing Pitch Phrases--24 items; Singing Rhythm Phrases--20 items; Singing Melodic Phrases--20 items; total test--76 items)

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#1	SI	.8315	.8418	.8408	.8445	.33
N=62	SP	.9223	.9333	.9570	.9497	
	SR	.8457	.7745	.7475	.7470	
	SM	.8992	.9224	.8888	.8943	
	CSS76	.9439	.9590	.9678	.9615	
UG#2	SI	.7865	.8288	.8432	.8022	.58
N=15	SP	.9209	.9299	.9320	.9436	
	SR	.7272	.7963	.7006	.7126	
	SM	.8921	.9185	.8923	.9127	
	CSS76	.9349	.9522	.9513	.9605	

Sample-N	Test	KRAP	KRIC	ORDSH	SELSH	"r".01*
UG#3	SI	.8456	.8231	.8874	.8977	.36
N=41	SP	.9148	.9236	.9663	.9244	
	SR	.7503	.7953	.7424	.7419	
	SM	.9110	.9373	.8847	.9229	
	CSS76	.9393	.9518	.9545	.9663	
UG#4	SI	.7854	.8286	.8064	.8595	.44
N=29	SP	.8983	.9100	.8951	.9231	
	SR	.8485	.8243	.8881	.8634	
	SM	.9056	.8852	.9432	.9192	
	CSS76	.9449	.9559	.9287	.9752	
UG#1-4	SI	.8348	.8446	.8502	.8661	.21
N=147	SP	.9203	.9280	.9445	.9397	
	SR	.8724	.8697	.8260	.8565	
	SM	.9094	.9205	.9061	.9080	
	CSS76	.9495	.9598	.9578	.9674	
UG#8	SI	.8551	.8685	.9003	.8497	.58
N=15	SP	.8698	.9010	.9299	.9212	
	SR	.7340	.7152	.7968	.5968	
	SM	.8313	.8655	.9104	.8704	
	CSS76	.9228	.9452	.9619	.9421	
G#1	SI	.8324	.7950	.7834	.7834	.26
N=97	SP	.9042	.9166	.9230	.9281	
	SR	.8810	.8644	.8330	.8757	
	SM	.8969	.9139	.9138	.9020	
	CSS76	.9463	.9562	.9515	.9624	
G#2	SI	.8573	.8552	.8488	.8558	.42
N=36	SP	.8809	.9027	.9066	.8920	
	SR	.7384	.7252	.7208	.7336	
	SM	.8602	.8971	.9176	.9033	
	CSS76	.9362	.9532	.9482	.9608	

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

TABLE XV --

4. THE WITH-IN TEST CORRELATIONS FOR THE SAMPLES LISTED, SEPARATELY, AND IN CERTAIN COMBINATIONS

Sample-N	SI/ SP	SI/ SR	SI/ SM	SI/ CSS	SP/ SR	SP/ SM	SP/ CSS	SR/ SM	SR/ CSS	SM/ CSS	"r" .01*
UG#1-62	.91	.28	.86	.91	.34	.89	.95	.43	.55	.95	.32
UG#2-15	.90	.22	.91	.90	.37	.96	.97	.34	.53	.96	.58
UG#3-41	.88	.19	.80	.89	.17	.92	.95	.20	.41	.94	.41
UG#4-29	.81	.27	.76	.85	.25	.89	.93	.28	.53	.92	.44
UG#1-4	.89	.28	.84	.90	.28	.91	.95	.32	.52	.94	.21
N=147											
UG#8-15	.83	.32	.46	.82	.46	.55	.91	.37	.66	.76	.58
G#1-97	.83	.25	.73	.81	.35	.85	.92	.52	.63	.94	.26
G#2-36	.89	.47	.87	.93	.41	.92	.96	.41	.59	.95	.42

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

THE PERTINENT TEST PAIRS

1. The PRM78 Dictation Test -- pre-test and re-test

TABLE XVI -- THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN CORRELATED MEANS (Reliabilities; Mean gain; "t" ratio; and interpretation)

Sample-N	Test	Rel-pre	Rel-re	Mean Gain	SE/DM	"t" ratio	Remarks
UG#1-62	Pitch--- 30 items	.86	.87	2.69	.66	4.5	Very Sig.
	Rhythm-- 24 items	.74	.73	2.16	.42	5.1	Very Sig.
	Melody-- 24 items	.73	.69	1.84	.43	4.3	Very Sig.
	PRM78	.90	.90	6.69	.97	6.9	Very Sig.
UG#2-15	Pitch--- 30 items	.74	.88	3.93	1.59	2.5	Nearly .01 level
	Rhythm-- 24 items	.70	.66	1.73	.56	3.1	Very Sig.
	Melody-- 24 items	.56	.74	.97	.76	1.3	Not Sig.
	PRM78	.84	.91	6.53	2.01	3.3	Very Sig.

Sample-N	Test	Rel-pre	Rel-re	Mean Gain	SE/DM	"t" ratio	Remarks
UG#3-41	Pitch---30 items	.87	.89	2.24	.58	3.8	Very Sig.
	Rhythm--24 items	.61	.57	2.68	.38	7.1	Very Sig.
	Melody--24 items	.72	.69	1.20	.55	2.2	Above .05 level
	PRM78	.88	.88	6.12	.96	6.4	Very Sig.
UG#4-29	Pitch---30 items	.78	.89	2.62	.88	3.0	Very Sig.
	Rhythm--24 items	.66	.57	2.93	.59	5.0	Very Sig.
	Melody--24 items	.59	.69	2.59	.54	4.8	Very Sig.
	PRM78	.81	.89	7.14	1.58	4.5	Very Sig.
UG#1-4 N=147	Pitch---30 items	.85	.89	2.68	.34	7.9	Very Sig.
	Rhythm--24 items	.71	.68	2.42	.26	9.3	Very Sig.
	Melody--24 items	.71	.73	1.71	.29	5.9	Very Sig.
	PRM78	.89	.90	6.80	.62	10.9	Very Sig.
UG#5-30	Pitch---30 items	.71	.64	.37	.51	.9	Not Sig.
	Rhythm--24 items	.64	.60	2.67	.68	4.0	Very Sig.
	Melody--24 items	.61	.63	1.10	.61	1.8	Near .05 level
	PRM78	.80	.83	4.14	1.41	2.9	Very Sig.
UG#6-51	Pitch---30 items	.87	.86	.10	.42	.2	Not Sig.
	Rhythm--24 items	.52	.70	1.80	.51	3.5	Very Sig.
	Melody--24 items	.75	.65	.63	.49	1.2	Not Sig.
	PRM78	.87	.88	2.53	.79	3.2	Very Sig.

Sample-N	Test	Rel-pre	Rel-re	Mean Gain	SE/DM	"t" ratio	Remarks
UG#1-6	Pitch---	.84	.88	1.63	.26	6.3	Very Sig.
N=247	30 items						
	Rhythm--	.68	.67	2.24	.25	9.0	Very Sig.
	24 items						
	Melody--	.71	.71	1.16	.23	5.0	Very Sig.
	24 items						
	PRM78	.88	.89	5.03	.50	10.0	Very Sig.

2. The PRM78 Dictation Re-test and the CSS76 Criterion Sightsinging Test

a. TABLE XVII -- THE CORRELATION BETWEEN MUSIC DICTATION AND SIGHTSINGING

		PRM78-re--		CSS76-----		Correlation	"r" .01*
Sample-N	Test	Rel'y	SE	Rel'y	SE		
UG#1-62	Pitch	.87	.031	.96	.010	.81	.32
	Rhythm	.74	.058	.96	.010	.28	
	Melodic	.73	.060	.96	.010	.76	
	PRM78	.90	.024	.96	.010	.76	
UG#2-15	Pitch	.74	.120	.95	.027	.88	.58
	Rhythm	.70	.136	.95	.027	.37	
	Melodic	.56	.184	.95	.027	.71	
	PRM78	.84	.078	.95	.027	.91	
UG#3-41	Pitch	.87	.038	.95	.016	.85	.41
	Rhythm	.61	.100	.95	.016	.20	
	Melodic	.72	.076	.95	.016	.70	
	PRM78	.88	.036	.95	.016	.84	
UG#4-29	Pitch	.78	.074	.95	.019	.66	.44
	Rhythm	.66	.106	.95	.019	.46	
	Melodic	.59	.123	.95	.019	.73	
	PRM78	.81	.064	.95	.019	.81	
UG#1-4	Pitch	.85	.023	.95	.008	.80	.21
N=147	Rhythm	.71	.044	.95	.008	.35	
	Melodic	.71	.044	.95	.008	.73	
	PRM78	.89	.017	.95	.008	.83	

Sample-N	Test	PRM78-re-- Rel'y	SE	CSS76----- Rel'y	SE	Corre-lation	"r".01*
G#1-97	Pitch	.85	.028	.95	.012	.86	.26
	Rhythm	.61	.064	.95	.012	.39	
	Melodic	.75	.045	.95	.012	.74	
	PRM78	.91	.018	.95	.012	.87	
G#2-36	Pitch	.89	.038	.95	.016	.84	.42
	Rhythm	.73	.079	.95	.016	.48	
	Melodic	.77	.069	.95	.016	.73	
	PRM78	.91	.029	.95	.016	.88	
UG#8-15	Pitch	.84	.079	.94	.031	.65	.58
	Rhythm	.50	.201	.94	.031	.17	
	Melodic	.41	.222	.94	.031	.25	
	PRM78	.73	.124	.94	.031	.64	

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

b. TABLE XVIII -- INTER-CORRELATIONS BETWEEN PRM78 DICTATION RE-TEST AND CSS76 CRITERION SIGHTSINGING TEST -- FOUR FIRST-YEAR SAMPLES, N=147

	PRM78---------- P. Item	R. Item	M. Item	PRM 78	CSS76---------- Singing---------- Int's	P. Ph.	R. Ph.	M. Ph.	CSS 76	"r".01*
Pitch Items	1.00	.39	.74	.92	.73	.81	.27	.82	.80	.21
Rhythm Items	.39	1.00	.46	.66	.27	.30	.59	.34	.44	.21
Melody Items	.74	.46	1.00	.88	.71	.72	.40	.73	.77	.21
PRM78	.92	.66	.88	1.00	.72	.78	.46	.80	.83	.21
Inter-vals	.73	.27	.71	.72	1.00	.89	.28	.84	.90	.21
Pitch Phs.	.81	.30	.72	.78	.89	1.00	.28	.91	.95	.21
Rhythm Phs.	.27	.59	.40	.46	.28	.28	1.00	.32	.54	.21

	PRM78				CSS76 Singing					
	P. Item	R. Item	M. Item	PRM 78	Int's	P. Ph.	R. Ph.	M. Ph.	CSS 76	"r".01*
Melody	.82	.34	.73	.80	.84	.91	.32	1.00	.94	.21
Phs.										
CSS76	.80	.44	.77	.83	.90	.95	.54	.94	1.00	.21

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

3. PRM78 Dictation pre-test and CSS76 Criterion Sight-singing Test

a. TABLE XIX -- CORRELATION OF MUSIC DICTATION (early Fall 1965) WITH SIGHTSINGING (Spring of 1966) -- FOUR FIRST-YEAR SAMPLES, N=147

Sample-N	Test	PRM78-re-- Rel'y	SE	CSS76----- Rel'y	SE	Correlation	"r".01*
UG#1-62	Pitch	.86	.031	.96	.010	.81	.32
	Rhythm	.66	.058	.96	.010	.28	
	Melodic	.74	.060	.96	.010	.72	
	PRM78	.90	.024	.96	.010	.76	
UG#2-15	Pitch	.74	.120	.95	.027	.88	.58
	Rhythm	.68	.136	.95	.027	.37	
	Molodic	.57	.184	.95	.027	.71	
	PRM78	.83	.076	.95	.027	.91	
UG#3-41	Pitch	.87	.038	.95	.016	.85	.41
	Rhythm	.61	.100	.95	.016	.20	
	Molodic	.69	.082	.95	.016	.70	
	PRM78	.88	.036	.95	.016	.84	
UG#4-29	Pitch	.78	.074	.95	.019	.66	.44
	Rhythm	.66	.106	.95	.019	.46	
	Molodic	.59	.123	.95	.019	.73	
	PRM78	.82	.064	.95	.019	.81	
UG#1-4	Pitch	.85	.023	.96	.008	.80	.21
N=147	Rhythm	.71	.044	.96	.008	.35	
	Molodic	.71	.044	.96	.008	.73	
	PRM78	.89	.017	.96	.008	.82	

* Assuming the null hypothesis of no correlation using the appropriate degrees of freedom.

b. An interesting group of correlations are those relating the PRM78 Dictation pre-test, the PRM78 Dictation re-test and the CSS76 Criterion Sightsinging Test from the data derived from these four schools.

		Correlations----------		
Sample-N	Test	PRM78-pre	PRM78-re	CSS76
UG#1-4	PRM78-pre	---	.83	.82
N=142	PRM78-re	.83	---	.83
	CSS76	.82	.83	---

TABLE XX --

c. INTER-CORRELATIONS BETWEEN PRM78 DICTATION RE-TEST AND CSS76 CRITERION SIGHTSINGING TEST -- FOUR FIRST-YEAR SAMPLES, N=147

	PRM78-pre---------				CSS76----------				
					Singing----------				
Sample-N	P. Item	R. Item	M. Item	PRM 78	Int's	P. Ph.	R. Ph.	M. Ph.	CSS 76
Pitch Items	1.00	.28	.71	.89	.70	.81	.26	.84	.80
Rhythm Items	.28	1.00	.43	.62	.22	.23	.52	.25	.35
Melodic Items	.71	.43	1.00	.88	.63	.64	.47	.71	.73
PRM78	.89	.62	.88	1.00	.68	.75	.47	.80	.82
Intervals	.70	.22	.63	.68	1.00	.89	.28	.84	.90
Pitch Ph.	.81	.23	.64	.75	.89	1.00	.28	.91	.95
Rhythm Ph.	.26	.52	.47	.47	.28	.28	1.00	.32	.54
Melodic Ph.	.84	.25	.71	.80	.84	.91	.32	1.00	.94
CSS76	.80	.35	.73	.82	.90	.95	.54	.94	1.00

SECONDARY RELATIONSHIPS

The relationships between the PRM78 Dictation Test and the CSS76 Criterion Sightsinging Test with three other selected conditions was probed through the data

accumulated. Analyses of the test data relative to the degree sought, the area of the major instrument, and training on the major instrument are shown in several tables following. Each table includes the "F" variance ratio and when this is significant the pertinent "t" ratios are listed to pinpoint the areas of significant difference. In some instances a near-significant "F" ratio is accompanied by one or more "t" ratios at or near the five-percent level. When the "F" ratios are non-significant the "t" ratios are either omitted, or, if their trend is of interest they are listed.

1. TABLE XXI -- THE RELATIONSHIP BETWEEN TYPE OF DEGREE SOUGHT AND ACHIEVEMENT IN MUSIC DICTATION AND SIGHTSINGING (*=.05 and **=.01 levels for "F" and "t")

a. Four first-year samples, N=147

Test Section	F Ratio		BA(36)	BM(80)	NON(26)
PRM78-pre	.32	BA-----	---	---	---
		BM-----	---	---	---
		NON----	---	---	---
CSS76	.89	BA-----	---	---	---
		BM-----	---	---	---
		NON----	---	---	---

b. Seven First-year samples, N=332

Test Section	F Ratio		BA(52)	BM(332)	NON(56)
PRM78-re Pitch	1.25	BA-----	---	---	---
		BM-----	---	---	---
		NON----	---	---	---
Rhythm	2.78*	BA-----	---	.63	2.39*
		BM-----	- .63	---	2.24*
		NON----	-2.39*	-2.42*	---
Melodic	1.43	BA-----	---	---	---
		BM-----	---	---	---
		NON----	---	---	---

Test Section	F Ratio		BA(52)	BM(332)	NON(56)
PRM78	2.26*	BA-----	---	1.51	2.21*
		BM-----	-1.51	---	1.29
		NON----	-2.21	-1.29	---

c. Four first-year and two graduate samples with N's as listed.

Test Section	F Ratio		UG#1-4 N=147	M.A. N=94	Ph.D. N=37
PRM78-pre	The "F" ratios were not computed for these comparisons.		No comparisons possible on pre-test.		
PRM78-re	Not computed	UG#1-4	.00	-1.67	-2.86**
		M.A.	1.67	.00	-1.99*
		Ph.D.	2.86**	1.99*	.00
CSS76	---	UG#1-4	.00	-3.48**	-6.40**
		M.A.	3.48**	.00	-2.75**
		Ph.D.	6.40**	2.75**	.00

2. TABLE XXII -- THE RELATIONSHIP BETWEEN THE AREA OF INSTRUMENTAL MAJOR AND ACHIEVEMENT IN MUSIC DICTATION AND SIGHTSINGING (*=.05 and **=0.1 for "F" and "t" ratios)

St-Strings; Kd-Keyboard; Ww-Woodwind; Br-Brass; Vo-Vocal; Pe-Percussion

Sample-N	Test		St	Kd	Ww	Br	Vo	Pe
a. UG#1-4 N=142	PRM78-pre	St	.00	2.02*	3.19**	2.29*	3.50**	2.04*
		Kd	-2.00*	.00	2.09*	.53	2.63**	.86
	F Ratio	Ww	-3.19**	-2.09*	.00	-1.57	.44	-.13
	3.45**	Br	-2.29*	-.53	1.57	.00	2.09*	.62
		Vo	-3.50**	-2.63**	-.44	-2.09*	.00	-.35
		Pe	-2.04*	-.86	.13	-.62	.35	.00
	PRM78-re	St	.00	2.11*	3.51**	2.80**	3.19**	1.69
	F=3.16**	Kd	-2.11*	.00	2.45*	1.22	1.92	.36
		Ww	-3.51**	-2.45*	.00	-1.29	-.59	-.79
		Br	-2.80**	-1.22	1.29	.00	.71	-.18
		Vo	-3.19**	-1.92	.59	-.71	.00	-.51
		Pe	-1.69	-.36	.79	-1.53	.51	.00

Sample-N	Test		St	Kd	Ww	Br	Vo	Pe
	CSS76	St	.00	1.60	3.38**	2.21*	2.36*	1.86
	F=3.51**	Kd	-1.60	.00	3.08**	1.09	1.34	.89
		Ww	-3.38**	-3.01**	.00	-2.03*	-1.73	- .57
		Br	-2.21*	-1.09	2.03*	.00	.28	.40
		Vo	-2.36*	-1.34	1.73	- .28	.00	.27
		Pe	-1.86	- .89	.57	- .40	- .27	.00

Because of a rather large F ratio in the rhythm section of the next sample (UG#1-7; N=330), each of the sections of the PRM78 Dictation re-test are detailed following as the next part of this same table.

b.	PRM78-re	St	.00	1.27	3.16**	2.43*	2.18*	1.51
UG#1-7	Pitch	Kd	-1.27	.00	3.40**	2.17*	1.69	.81
N=330	section	Ww	-3.16**	-3.40**	.00	-1.35	-1.63	-.76
	F=3.52**	Br	-2.43*	-2.17	1.35	.00	- .36	-.13
		Vo	-2.18	-1.69	1.63	.36	.00	.04
		Pe	-1.51	- .81	.73	.13	- .04	.00
	Rhythm	St	.00	1.16	- .25	- .27	1.85	-2.03*
	section	Kd	-1.16	.00	-2.42*	-2.63**	1.30	-3.37**
	F=5.84**	Ww	.25	2.42*	.00	- .02	3.44**	-2.19*
		Br	.27	2.63**	.02	.00	3.69**	-2.23*
		Vo	-1.85	-1.30	-3.44**	-3.69**	.00	-3.90**
		Pe	2.03*	3.37**	2.19*	2.21*	3.90	.00
	Melody	St	.00	1.28	2.29*	2.25*	1.94	1.70
	section	Kd	-1.28	.00	1.85	1.83	1.24	1.04
	F=1.84	Ww	-2.29*	-1.85	.00	- .15	- .59	.17
		Br	-2.25*	-1.83	.15	.00	- .48	.25
		Vo	-1.94	-1.24	.59	.48	.00	.47
		Pe	-1.70	-1.04	- .17	- .25	- .47	.00
	whole	St	.00	1.47	2.45*	2.03*	2.41*	.82
	test	Kd	-1.47	.00	1.80	1.05	1.74	- .18
	F=1.90	Ww	-2.45*	-1.80	.00	- .81	- .10	-1.01
		Br	-2.03*	-1.05	.81	.00	.72	- .64
		Vo	-2.41*	-1.74	.10	- .72	.00	- .96
		Pe	- .82	.18	1.01	.64	.96	.00

Sample-N	Test		St	Kd	Ww	Br	Vo	Pe
c.	PRM78-re	St	.00	1.17	1.09	1.58	1.19	.56
SO&Gr		Kd	-1.17	.00	.03	.69	.04	.03
N=148		Ww	-1.09	- .03	.00	.55	.00	.02
		Br	-1.58	- .69	- .55	.00	- .66	- .19
		Vo	-1.19	- .04	.00	.66	.00	.02
		Pe	- .56	- .03	- .02	.19	- .02	.00
	CSS76	St	.00	.86	1.42	1.68	1.27	1.22
	F=.88	Kd	- .86	.00	.92	1.31	.67	.89
		Ww	-1.42	- .92	.00	.24	- .37	.55
		Br	-1.68	-1.31	- .24	.00	- .69	.46
		Vo	-1.27	- .67	.37	.69	.00	.69
		Pe	-1.22	- .09	- .55	- .46	- .69	.00

3. TABLE XXIII -- THE RELATIONSHIP BETWEEN TRAINING ON MAJOR INSTRUMENT AND ACHIEVEMENT IN MUSIC DICTATION AND SIGHTSINGING (*=.05 and **=0.1 for "F" and "t" ratios)

a. Four first-year samples, N=145

Test	F Ratio		Years of Training 9,up	6-8	3-5	2,less
PRM78-pre	3.03*		N=48	N=43	N=27	N=27
		9,up	.00	1.25	1.57	2.96**
		6-8	-1.25	.00	.48	1.84
		3-5	-1.57	- .48	.00	1.23
		2,less	-2.96**	-1.84	-1.23	.00
PRM78-re	2.78*	9,up	.00	.01	1.81	2.28*
		6-8	- .01	.00	1.77	2.23*
		3-5	-1.81	-1.77	.00	.41
		2,less	-2.28*	-2.23	- .41	.00
CSS76	1.21*	9,up	.00	.99	1.06	1.85
		6-8	- .99	.00	.19	.97
		3-5	-1.06	- .19	.00	.70
		2,less	-1.85	- .97	- .70	.00

b. Seven first-year samples, N=337

PRM78-re	2.81*		N=105	N=97	N=73	N=62
		9,up	.00	- .02	2.03*	.78
		6-8	.02	.00	1.99*	.77
		3-5	-2.03	-1.99	.00	-1.07
		2,less	- .78	- .77	1.07	.00

c. One sophomore and two graduate samples, N=148 ("F" at .05=2.68, and "F" at .01=3.93(

There seems to be no clear pattern relating training on the major instrument with achievement in music dictation and sightsinging. The "F" ratios for variables including the rhythm factor seem to be the largest, but the few significant "t" ratios are apparently scattered at random. The "F" ratios for each section of the dictation and sightsinging test are given.

Pitch Section (PRM78)	.49
Rhythm Section (PRM78)	1.63
Melodic Section (PRM78)	.73
PRM78	.96
Intervals (CSS76)	.32
Pitch Phrases (CSS76)	.25
Rhythm Phrases (CSS76)	1.67
Melodic Phrases (CSS76)	1.55
CSS76	.95

DISCUSSION

Prior to doing the testing accomplished in this project it was known that students beginning music study at various schools had at least two factors in common, namely, the desire and some background for doing serious study in music, as evidenced by enrollment in first-year music theory and eartraining, together with other classwork. Bartlett's test of the homogeneity of variance gave clear indication that all of the six first-year samples that took the PRM78 Dictation pre-test and the thirteen that took the PRM78 Dictation re-test and the seven samples that took the CSS76 Criterion Sightsinging Test--a total of thirteen different samples from eight schools--were from the same parent population of musicians, perhaps further identifiable as those specializing in the field of music. It was a revelation, though not wholly unexpected as can be seen from the fourth hypothesis, that the three graduate student samples were shown to be members of the same population as the undergraduate students, comprising mainly first-year music students. The largest chi-square obtained was that resulting from

combining all thirteen samples and its value was 18.5. However, since a chi-square of 21.5 is necessary at the five-percent level and one of 26.5 at the one-percent, it is obvious that a chi-square of 18.5 is somewhat short of the five-percent level of confidence.

The six first-year samples differed in mean attainment on the dictation test, both the pre-test and the re-test. However, each sample achieved statistically significant improvement during the 7-1/2 month period of time separating the pre-test and the re-test. This clearly attests to the affirmation of the third hypothesis concerning the expectation that significant growth in dictation achievement would take place during the first-year of eartraining study. This confirmation of significant growth in dictation achievement does not suggest that we should be satisfied. Rather, it should make us still more curious about how much attainment can be accomplished maximally, how much this level differs from what we routinely accomplish, and how rapidly we can accomplish significant gains in this area.

It is necessary that samples be homogeneous in variance in order to make valid comparisons; it is equally necessary that the test instruments used be reliable statistically. Several standard methods of measuring test reliability are available.

1. Test, followed soon by a re-test. This method was ruled out because the dictation test was to be repeated following a school year of eartraining in order to measure gain in dictation achievement. However, somewhat unexpectedly, the correlation after 7-1/2 months of eartraining study was still a very significant .83, measured from a combined sample of 147 students from 4 schools. While this correlation can serve as one important statistic for predicting end-of-year achievement from entrance scores, it simultaneously does constitute a significant measure of reliability, indicating that even after a period of training the pre- and re- rank-orders had much in common. This leaves a hint that any eartraining program followed has a real up-hill task to accomplish.

2. Equivalent tests. This method was ruled out because an equivalent dictation test was not available for use in this project.

3. Therefore intensive study of the two separate instruments was done using the four measures of reliability discussed previously.

It was encouraging that the four measures of reliability computed for each sample on each test taken by that sample turned out to be quite consistent. In 87 of the 111 sets of reliability coefficients computed for the two tests involving the 13 samples, the range of the four measurements for a given test did not exceed .10. The general level of reliability for the PRM78 Dictation Test was about .88 (pre-test and re-test differed overall only about .01) for undergraduates and about .90 for graduate students. The level for the dictation test sections were about .86 for pitch, about .68 for rhythm, and about .73 for melody. The general level for the CSS76 Criterion Sightsinging Test was about .95 for all seven samples, with section reliabilities of about .84 for singing intervals, about .93 for singing pitch phrases, about .82 for singing rhythm phrases, and about .92 for singing melodic phrases. All of these levels were considered satisfactory at this stage of our investigations in eartraining, but it is our desire to do future research to attempt to raise the reliability of each section up to the .90 level. Parenthetically, since each test's reliability was listed in terms of the four separate measures taken, the coefficients given here represent an approximate arithmetical average. Although correlation coefficients are not normally averaged except through the use of Fisher's "z", this method was not employed in this project for the following reasons: 1) the coefficients were almost always clustered within a small range, 2) use of the "z" intermediary technique tended to raise slightly the averages given in this report, and 3) this investigator was willing to accept the slight underestimates involved in using arithmetic averages.

After samples were found to be homogeneous and the testing instruments quite reliable, it became

logical to assess the possible intercorrelations. The first computer program used readily worked out the four-way reliabilities mentioned above, and the second one gave all within-test intercorrelations as well as the between-test intercorrelations for each section as well as for the total tests in addition to accomplishing variance analysis. Whereas intercorrelations within a test are indications of the relation of each section to the whole test, all intercorrelations between different tests are in actuality validity measures, that is, they show the degree of relationship between a given section or an entire first test and a given section or the whole of a second test. In this phase the PRM78 Dictation Test and its sections were correlated with the CSS76 Criterion Sightsinging Test and its sections. The general range and the approximate means of validity coefficients deriving from this investigation are cited following in support of the truth of the first and second hypothesis concerning the correlations between the PRM78 Dictation and the CSS76 Criterion Sightsinging test.

1. Pitch Section (PRM78) with CSS76 Test -- two-thirds of the coefficients fall between .80 and .86, with a mean of about .82.

2. Rhythm Section (PRM78) with CSS76 Test -- two-thirds of the coefficients fall between .35 and .48, with a mean of about .43.

3. Melodic Section (PRM78) with CSS76 Test -- two-thirds of the coefficients fall between .70 to .78, with a mean of about .75.

4. PRM78 Dictation Test with CSS76 Test -- two-thirds of the coefficients fall between .76 and .91, with a mean of about .85.

In summarizing the growth of achievement in music dictation premised in the third hypothesis, 17 out of 24 "t" ratios possible from the six schools concerned were found highly significant at well above the one-percent level with another almost at the same level and two more at the five-percent level. These "t"

ratios were derived from calculating the significance of the difference between correlated means based upon the pre-test and re-test scores of the PRM78 Dictation Test. The "t" scores from the cumulative sample were highly significant in asserting further that positive change had been accomplished.

The fourth hypothesis premised that the difference between graduate and undergraduate students were favorable to the former but not necessarily significant. The comparisons made in dictation and sightsinging involved four end-of-first-year samples and two graduate samples and these provided statistics of considerable interest. In the dictation test those seeking the masters' degree achieved only a favorable, but not a significant difference, with a lone margin of significant difference in the pitch section. However, their margin was significant in the sightsinging and all of its sections. Comparison with those seeking the doctorate provided the best refutation of the fourth hypothesis. Even though doctoral candidates were very significantly above undergraduates in both total tests and the masters' candidates in the sightinging test, the rhythm section of the dictation test proved to be non-significantly different among all tested. Perhaps a full explanation is not forthcoming at this moment, but this investigator feels that perhaps most schools do not work as ambitiously at acquiring proficiency in rhythm dictation as they do in pitch dictation. Summarizing, graduate students seemed to excell in pitch concepts and undergraduates posed their greatest challenge in those areas including the rhythm factor. Therefore, in the opinion of this investigator, the refutation of the fourth hypothesis regarding graduate and undergraduate differences was somewhat ambiguous and hence not without equivocation.

On the undergraduate level there was little difference between those seeking the "BA" or the "BM" degrees, except that both of these surpassed in rhythm dictation (at the five-percent level) those listed as "non-majors", and that those seeking the "BA" differed at the five-percent level on the total dictation test from those seeking the "BM". However, at the graduate

level, those seeking the doctorate were quite close to significant in differing from masters' candidates in music dictation, who in turn held a similar relationship with the first-year samples. In the total sightsinging test doctoral students were significantly higher than masters candidates who were significantly higher than the undergraduate samples. Possibly the proficiency in sightsinging, whose serious development most often does not begin until undergraduate days, is quite related to the amount of college music experience of students.

Analysis of the data from four first-year samples showed that a small group of "string majors" made significantly higher scores than any of the other instrumental groups on the PRM78 Dictation pre-test and this was sustained on both the PRM78 Dictation re-test, and also on the CSS76 Criterion Sightsinging Test. To a lesser degree and not always significantly in relation to those below them this also applied to those classified as "keyboard majors." The "woodwind" and "brass" majors formed a third group much by themselves with the vocal and small group of percussionists at the lower level of achievement, with the latter strong, of course, at rhythm. In the group of 7 first-year groups on the PRM78 Dictation re-test much the same pattern prevailed except that the vocal groups encouragingly moved near the woodwind and brass in other than the area of rhythm. Considering graduate students, apparently cumulative college experience is a moderating factor on the effect of the instrumental major area, since even though some trends were present, no significant "t" ratio differences were found.

In general, the extent of instrumental training and achievement in dictation and sightsinging showed a common linear trend toward significance, but this was never actually significant except in the case of the two extremes of the categories, those with more than nine years of training and those with two or less years of training. This tendency was even less pronounced and in fact, hardly discernible, on the graduate level.

CONCLUSIONS

All conclusions, implications, and recommendations made are limited to the experience of this investigator gained from or corrobcrated through the data and statistics derived from this project.

1. The 13 samples participating in this project--7 first-year undergraduate samples, 3 second and third year samples, and 3 graudate samples--proved to be homogeneous in variance through use of the Bartlett Test.

IMPLICATIONS--This possibly opens the door to the broad application of research techniques to studying various aspects of eartraining over the entire undergraduate and graduate span of students.

2. The CSS76 Criterion Sightsinging Test proved to be a highly reliable test (ca. .95) on both the undergraduate and the graduate levels.

RECOMMENDATIONS--1. Further study might be done to compare the reliabilities among several test administrators. This was not done in this project, since the overall reliabilities achieved were invariably above the .90 level.

2. Further study should be done to determine to what extend a few, selected, weaker items can be improved by alteration or replacement.

3. Further study should be done to determine whether the length is now optimal (ca. 30 minutes per each individual tested), or whether the length of test can be reduced without critical loss of reliability.

3. The PRM78 Dictation Test proved to be quite reliable, with a mean coefficient of about .88. In the opinion of this investigator the test is based upon a premise that is fundamentally tenable, namely, that a trained musician should be aware of how pitch, rhythm, and melodic sounds relate to the pertinent symbols of notation.

IMPLICATION--Since this test does not include harmonic or contrapuntal aural awareness in testing dictation achievement, it would be interesting to premise that achievement in these areas would correlate substantially with proficiency in the PRM78 Dictation Test.

RECOMMENDATIONS--1. Do exhaustive item analysis from every logical point of view, and on the basis of such further and related study change possibly four of the thirty items in the pitch section, perhaps up to one-half of the twenty-four items in the rhythm section, and up to one-third of the items in the melodic section.

2. Determine whether the use of both bass and treble clefs would be more logical and reliable than only that of treble clef as now is done.

4. The rhythm factor (and hence the melodic factor as well) is far more difficult to test reliably than is the pitch factor. The reasons for this being true are not fully clear to this investigator at this point, but several ideas are suggested.

IMPLICATIONS--There is a real possibility that the recorded test performance of the rhythm items constituted an important source of error variance. In this test the repetition of the item was a separate performance, which undoubtedly varied to some extent, introducing a certain degree of subjectivity into the item's recognition. But going further, perhaps any performance of a complicated rhythm by a human being rather than by a machine does become somewhat subjective in a subtle sense. An example would be the variety that could be found for the performance of the dotted eighth and sixteenth combination of notes. Perhaps machine performance of rhythm patterms is the most logical and objective manner in which to proceed for didactic purposes.

RECOMMENDATIONS--Considerable research is needed in the area of rhythm, covering specifically, item difficulty, the comparative reliability of items performed by machine and those humanly done, as well as basic research on student reactions and responsiveness to rhythm dictation done by these methods.

5. The PRM78 Dictation Test (reliability of ca. .88) correlated with the CSS76 Criterion Sightsinging Test (reliability of ca. .95) at a substantial level of about .85.

IMPLICATION--Highly reliable individual predictation requires a highly reliable test from which to predict a criterion with which it is known to correlate at a very high level.

RECOMMENDATION--Further research is necessary to raise the reliability of each test section, but especially rhythm and melody into the .90's, assuring that each added item, or changed item, itself correlates at a significant level with the CSS76 Criterion Sight-singing Test.

6. Ordinary methods of teaching eartraining are apparently relatively effective. However, all that was learned in this area from the project was that substantial improvement in dictation achievement was made. But, there was no indication available as to how this attained level related to the "maximum possible" mean attainments or the "maximum feasible" levels of mean attainment.

RECOMMENDATIONS--1. Further research comparing various methods of teaching eartraining and comparing their effectiveness should be done.

2. Further research directed toward deriving yet new methods for teaching eartraining should be pursued.

3. Specific research should be accomplished to investigate the "maximum possible" mean levels of achievement, as well as the more practical "maximum feasible" mean levels.

7. The differences between undergraduates and graduate students are not unambiguous at the masters' level but tend to become almost entirely so at the doctoral level when compared with the undergraduate level of achievement. The pitch variable was the main factor producing differences favoring the graduate students, whereas the rhythm variable was the main factor permitting undergraduate challenge of graduate attainment. However, the almost wholly favorable graduate position is sight-singing achievement suggests that perhaps the extent of college level experience is an important ingredient in this particular skill of the trained musician.

RECOMMENDATION--Further comparative research concerning undergraduate and graduate attainment in dictation and sightsinging is necessary, to supplement the beginnings made in this research project.

8. The selective process determining whether the BA or the BM degree is sought is perhaps more administrative than musical; there was only a slight, insignificant "t" ratio favoring the former degree as representative of higher achievement. But in several areas masters' candidates achieved higher than undergraduates and in most areas doctoral candidates surpassed both.

IMPLICATION--This suggests a confirmation of the logic of using highly selective and reliable tests for validating acceptance of candidacy for a degree on the graduate level.

9. The major instrument area and the extent of training are both important factors in music dictation and sight-singing achievement in the first-year class, and especially so when measured at the beginning of training. This influence seems to moderate in effect as the undergraduate training proceeds and leaves only non-significant traces at the graduate level.

IMPLICATION--Testing in the beginning of the school year can assist in selecting those with potential problems in eartraining.

RECOMMENDATIONS--Further research is needed to determine what if any relation there may be between the progress made in eartraining and the major instrument and the extent of training, both considered separately, as well as concurrently.

SUMMARY

Eartraining concerns developing proficiency in hearing pitch relations and rhythm relations, separately or in manifold combinations, and is generally thought of in terms of sightsinging and music dictation. Since sightsinging testing is time consuming though highly reliable, this investigation dealt with the development and evaluation of both a music dictation test (PRM78 Dictation Test) and a sightsinging test (CSS76 Criterion Sightsinging Test). It was hoped that the dictation test could eventually be developed to serve as an adequate replacement for the latter, designated as the criterion sightsinging test because sightsinging is generally accepted as a valid and substantial aspect of overall musicianship.

The objectives studied as hypotheses were that each section and the dictation test would be substantially correlated with the sightsinging test; that statistically significant growth in music dictation achievement would be attained in the first-year eartraining class; and finally that graduate music students achieve at a higher, but not necessarily significantly different level than first-year students attain upon completion of their first year of training. Secondary objectives were to study the relationships between degree sought, the area of the major instrument, and the extent of training on the major instrument with proficiency in music dictation and sightsinging.

The PRM78 Dictation Test given as a pre-test to the first-year students of six schools quite early in the school year, and the same test given as a re-test to the same students, as well as to seven other samples including one first-year, three above-first-year undergraduate, and three graduate samples yielded total test reliabilities ranging generally from about .82 to .91. The total test reliability for the samples considered cumulatively was about .88, with a pitch section reliability of .86, rhythm .68, and melody .73. The CSS76 Criterion Sightsinging Test produced total test reliabilities of about .95 in all samples, singly or cumulatively, with an interval section reliability of .84, pitch .93, rhythm .82, and melody .92.

The entire group of 13 samples--7 first-year, three second and third year, and 3 graduate--was proved to be homogeneous in variance through use of the Bartlett Test and this fact is pertinent in accepting the reliabilities mentioned above. since the mean coefficient levels given above are also applicable to the combined samples. With the homogeneity of variance affirmed, the first and second hypotheses concerning the correlation of the PRM78 Dictation Test--its sections and the test--with the CSS76 Criterion Sightsinging Test were upheld in the manner indicated following:

$r_{\text{Pitch section (PRM78) . CSS76}}$ = ca. .82,[1] generally; a very substantial level of correlation.

$r_{\text{Rhythm section (PRM78) . CSS76}}$ = ca. .43, generally, a moderate, but not substantial level of correlation. However, it is worthy to note that even this coefficient is above the one-percent level of significance necessary to refute the null hypothesis of no correlation between the two variables (ca. r = .24 is needed for 3 variables at 150 degrees of freedom).

$r_{\text{Melodic section (PRM78) . CSS76}}$ = ca. .75, generally; again a very substantial correlation.

$r_{\text{PRM78 . CSS76}}$ = ca. .85 generally; a very substantial level of correlation.

[1]--an explanation of the meaning of "average" coefficient of correlation was given on page 34.

The evaluation of the PRM Dictation pre-test with its repetition as a re-test after seven and one-half months of formal eartraining showed, overall, highly significant gains in all sections and the total test, with 17 out of the 24 possible means considerably above

the one-percent level, two more coefficients at five-percent and another nearly at the five-percent level of confidence. Furthermore, the cumulative sample showed highly significant changes in achievement level in all sections and in the total test, as revealed in "t" ratios ranging from 5.0 to 10.0. These results were considered strong affirmation of the third hypothesis concerning the significance of growth achieved during a year of formal eartraining. An interesting corollary to the reliability statistics mentioned earlier was that the correlation between the PRM78 Dictation pre-test and the re-test was .83. Several observations are made.

1. There was substantial correlation between the results achieved from giving the PRM78 Dictation Test in the Fall and again 7-1/2 months later. But, there was also a very significant gain made in mean achievement during this period of time. From these two facts one can infer that the rank-order of the students showed considerable stability on the two tests. Therefore, it would seem logical that meaningful prediction of achievement in the Spring can be made early in the Fall, assuming the availability of a highly reliable dictation test.

2. This substantial correlation is in essence a belated, but certainly a significant additional confirmation of the reliability of the PRM78 Dictation Test, established by a fifth method (four other methods were used previously), namely, that of test-retest.

3. However, it must be considered disturbing that almost a full school year's work in eartraining has not shifted scores and ranks sufficiently to severely disturb the test's reliability; as a pre-test its reliability was about .87, as a re-test about .88, and the pretest-retest correlation reliability is .83. It thus becomes obvious that continued and concentrated research is needed to uncover yet more effective teaching methods, but further, to evaluate the worth of such methods by extensive testing involving a broad cross-section of samples representative of the pertinent music population.

The fourth hypothesis premised that graduate student margins above undergraduate achievement in dictation and sightsinging would perhaps not be significant statistically. Comparing the four first-year classes at the end of the year with masters' degree candidates gives the latter only a non-significant margin in the dictation test, but this becomes quite significant in sightsinging. It must be noted, however, that the favorable margin for these graduate students was not significant in either rhythm dictation or rhythm sightsinging. Extending this comparison of first-year students to graduate doctoral candidates, the graduate margin becomes highly favorable in almost every aspect of dictation and sightsinging. The exception again is rhythm, but this lends credence to the point that graduate students seem to excell more often in pitch concepts and this is furthermore true regarding sightsinging. On the contrary, undergraduate students are able to pose their best challenge in the sections involving the rhythm factor.

For undergraduates there was indication at the five-percent level of confidence that non-music major students enrolled in first-year theory classes achieved less well than did those seeking the BA or the BM degrees and there was a non-significant tendency for those seeking the BM degree to achieve slightly lower scores than those seeking the BA degree. The difference in achievement in music dictation between those working on the doctoral and masters' degrees was significant at the five-percent level, favoring the doctoral candidates, and this widened when considering the undergraduates.

Discussing the four first-year samples that took the dictation pre-test and re-test and the sightsinging test, string majors tended to achieve above all other instrumental groups on the PRM78 Dictation Test. Pianists held a distinctly lesser second place, but there was less difference between brass, woodwind, vocal and percussion areas, except that the vocal group was low more often. In general this hierarchy was preserved at a lesser degree of significance in the re-test following 7-1/2 months of training, and in the sight-singing test the same general pattern was again present,

if not still more obvious. Dealing with the sample enlarged to include seven first-year classes in the re-test, the same tendency existed in the pitch section (PRM78), but in the rhythm section the percussion, brass and woodwind achieved the significant levels of difference. However, on melody the balance once more tipped to strings and pianists, moderating on the total test to favor only the strings. A study of 133 graduate and 15 sophomore students demonstrated clearly that "instrument category" was no longer as important for determining achievement level. There was a tendency for strings to achieve somewhat higher, but it was no longer significant in any of the possible comparisons. In general, these remarks were applicable to sightsinging as well as to dictation.

The last comparison involved assessing the influence of the extent of training on the major instrument upon proficiency in dictation and in sightsinging. Dealing with the same four first-year classes the pattern of "t" ratios achieved on the PRM78 Dictation pre-test clearly followed the pattern of extent of training; however, it required a comparison of the two extremes--those with nine or more years of training, and those with two or less years of training--to achieve significance at the one-percent level of confidence. Rather inconsistent with this neat parallel is the fact that the sample enlarged to the seven first-year samples did not follow this expected pattern as clearly. A possible explanation is that at least one and perhaps more of the schools involved are following a relatively rigorous program of eartraining, whose purpose would obviously be to improve the achievement of their students and this could tend to render the extent of background or training somewhat less of a factor for determining the extent of improvement in achievement.

In the opinion of this investigator the present research project has contributed to the credibility, clarification, or understanding of at least nine points concerning eartraining research.

1. The undergraduate and graduate samples used in this project were proved to be homogeneous in variance through

use of Bartlett's Test. This fact, affirming that all of these samples were part of the same parent population, is of considerable importance to future research.

2. Sightsinging can effectively and reliably be tested using an instrument whose evaluation in the hands of a competent test administrator is almost completely objective.

3. Highly reliable music dictation tests can be constructed. Regarding the instrument used in this project, selective efforts at improving perhaps four pitch items, about half of the rhythm items and about one-third of the melodic items would with little doubt permit the achievement of a reliability of near .95, comparable with that achieved in the sightsinging test.

4. The correlation between reliable dictation and sightsinging tests proved to be on a very substantial level. This gives credence to one hope of this investigator, namely, that achievement in sightsinging will be able to be predicted reliably using an objective dictation test. Research to increase the reliability of the current dictation test from .88 to .95 with a probability of achieving a dictation versus sightsinging correlation of at least .90 would boost the common variance between the two tests to a value of about 81% from its present level of about 72%.

5. The rhythm factor and also the melodic factor (primarily because it includes rhythm) are more difficult to test reliably than is the pitch factor. Further basic research in this area is essential to making additional progress.

6. The intercorrelation between achievement on pitch and rhythm concepts tends to fall most often around .40, a significant coefficient, but one that is indicative of the great difference between these two basic ingredients of music, the "vertical" and the "horizontal", whose common variance is perhaps about one-sixth of the total variance in each variable.

7. Ordinary methods of teaching eartraining are certainly not ineffective. However, additional research is needed to investigate just how much improvement is possible, how far this ultimate mean level surpasses current mean levels, and ultimately, how much improvement is feasible in relation to all factors involved.

8. The degree sought is not an important factor in determining achievement in dictation and sightsinging on the undergraduate level. However, on the graduate level it becomes increasingly important. Those seeking the masters' achieved better than undergraduates, not significantly in dictation, but definitely so in sightsinging. Continuing, those seeking the doctorate were in turn above the masters' candidates, non-significantly in dictation, but very significantly in sightsinging.

9. The major instrumental area is a significant factor on the undergraduate level and especially in the first-year class. This is also true regarding the extent of training. However, with training this difference moderates, this becoming obvious already in the retest at the end of the first year and finally becoming non-significant except for some trend at the graduate level.

REFERENCES

BOOKS

4. Edwards, Allen L., Experimental Design in Psychological Research, New York, Rinehart and Co., Inc., 1950.

6. Lundin, Robert W., An Objective Psychology of Music, New York, Ronalds Press Col, 1953.

10. Mursell, James L., The Psychology of Music, New York, W.W. Norton and Co., 1937.

16. Seashore, Carl E., The Psychology of Musical Talent, Boston, Silver, Burdett and Company, 1919.

18. Thompson, Randall, College Music, New York, The McMillan Co., 1935.

ARTICLES

2. Dean, Charles D., "Predicting Sightsinging Ability in Teacher Education," Journal of Educational Psycology, XXVIII, (November, 1937), 601-608.

3. Drake, R.M., "Four New Tests of Musical Talent", Journal of Psychology, XVII, (1933), 136, ff.

14. Revész, Geza, The Psychology of a Musical Progidy, 1920, cited in Holstrom, Lars-Gunnar, Music and Prognosis, Uppsala, Sweden, Almquist and Wiksells, 1953.

15. Salisbury, Frank S., and Smith, Harold B., "Prognosis of Sightsinging Ability of Normal School Students," Journal of Applied Psychology, XIII, (October, 1929), 425-439.

RESEARCH REPORTS

7. Madison, Thurber H., "Interval Discrimination as a Measure of Musical Talent," Archives of Psychology, XXXVIII, 268, (June, 1942), 1-99.

8. Marquis, James H., "A Study of Interval Problems in Sightsinging Performance with Consideration of the Effects of Context," Doctoral Dissertation, The University of Iowa, 1963.

9. Mosher, Raymond M., "A Study of the Group Method of Measurement of Sightsinging," Contributions to Education, No. 194, New York, Teachers College, Columbia University Press, 1925.

11. Ortmann, Otto, Problems in the Elements of Ear-Dictation, Baltimore, Maryland, Peabody Conservatory of Music, 1934.

12. Ottmann, Robert W., "A Statistical Investigation of the Influence of Selected Factors on the Skill of Sightsinging," Doctoral Dissertation, North Texas State University, 1956.

13. Poland, B.W., "An Investigation of Some Aural and Notational Elements in Music Theory," Doctoral Dissertation, Ohio State University, 1960.

17. Spohn, Charles L., "An Exploration in the Use of Recorded Teaching Material to Develop Aural Comprehension in College Music Classes", Doctoral Dissertation, Ohio State University, 1959.

_____, "An Evaluation of Two Methods Using Magnetic Tape Recordings for Programmed Instruction in the Elemental Materials of Music," Research Report, The Ohio State University Research Foundation, 1963.

TESTS

1. Aliferis, James, Aliferis Music Achievement Test, Minneapolis, University of Minnesota, 1954.

5. Gordon, Edwin, Gordon Measures of Musical Aptitude, New York, Houghton, Mifflin Co., 1964.

19. Wing, Herbert D., Musical Ability and Appreciation ....a New Series of Standardized Tests, The University of London, London, 1941.

NAME _____
Last

Over-all number in
TERP 65-66 _____
(leave blank)

THOSTENSON EARTRAINING RESEARCH PROJECT (1965-1966)

School _____

Over-all number within
participating school - _____
(leave blank)

Age _____ Sex _____

Intended major area

Music, BA ---------- _____

Music, BM or BME--- _____

Other music degree-- _____

Not in music-------- _____

Summary of Training in Music

Years, band _____ Years, choral _____ Years, orchestral _____

Instrument category	Which?		Years, lessons	Importance to you (1, 2, etc.)
- Keyboard----------	_____	_____	_____	_____
- Woodwind----------	_____		_____	_____
- Voice----------	_____		_____	_____
- Brass----------	_____		_____	_____
- Percussion----------	_____		_____	_____
- Strings----------	_____		_____	_____
- Other----------	_____		_____	_____

PRMA SCORES (leave blank)	P30	R24	M24	PRMA78	SSI	SSP	SSR	SSM	CSST
Fall 1965----------	___	___	___	_____					
Spring 1966----------	___	___	___	_____	___	___	___	___	____
Gain made----------	___	___	___	_____					

A-1

PITCH, RHYTHM, MELODIC ACHIEVEMENT DICTATION TEST

PITCH

NAME _____ #_____

ANSWER

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
A											
B											
C											
D											

NAME_____

NSWER

PITCH

A B C D

13.

14.

15.

16.

17.

18.

19.

20.

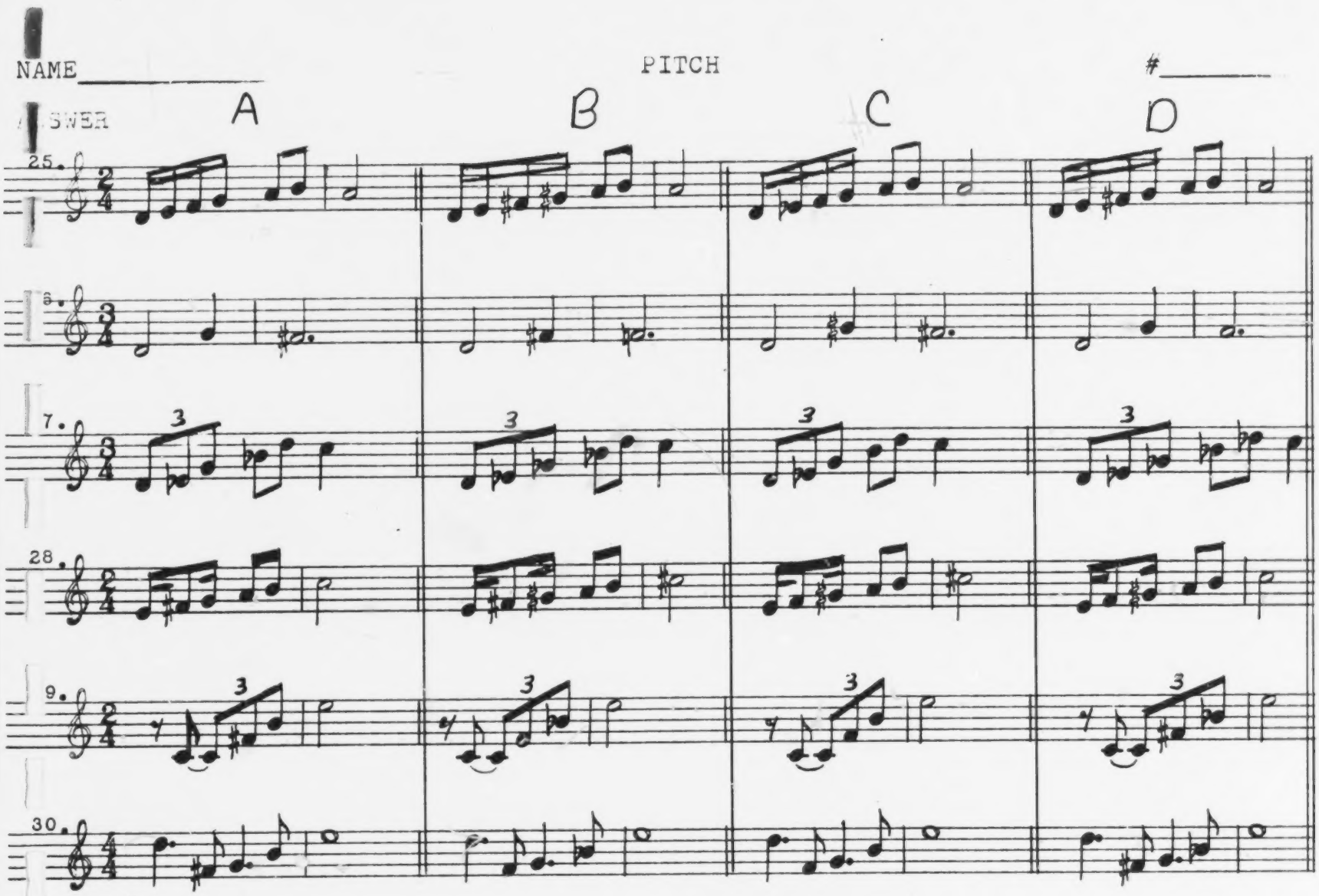
21.

22.

23.

24.

NAME_____ PITCH #_____

ANSWER A B C D

NAME _____ #_____

RHYTHM

ANSWER

A B C D

NAME _____ RHYTHM #_____

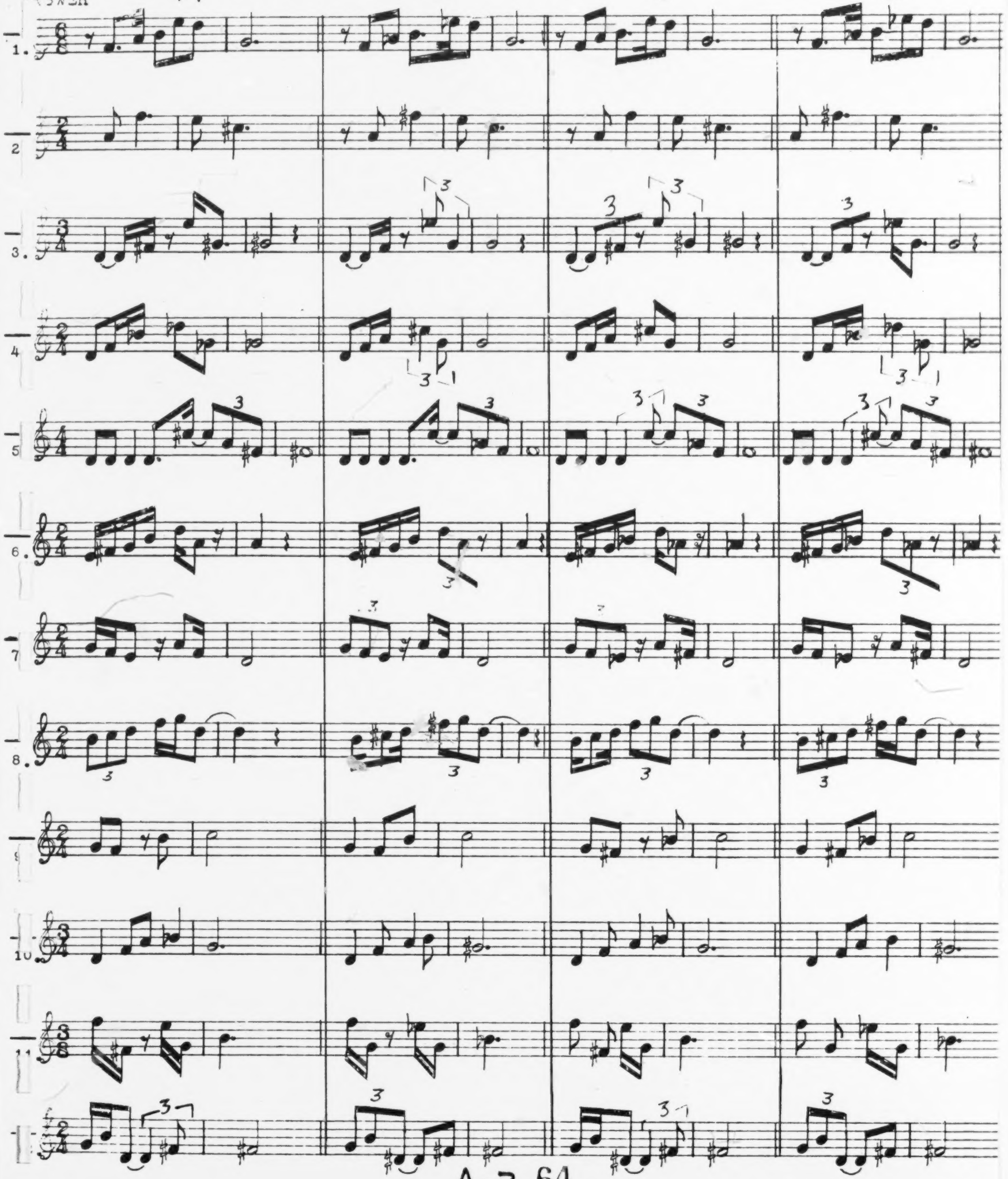
ISWER

A B C D

MELODY

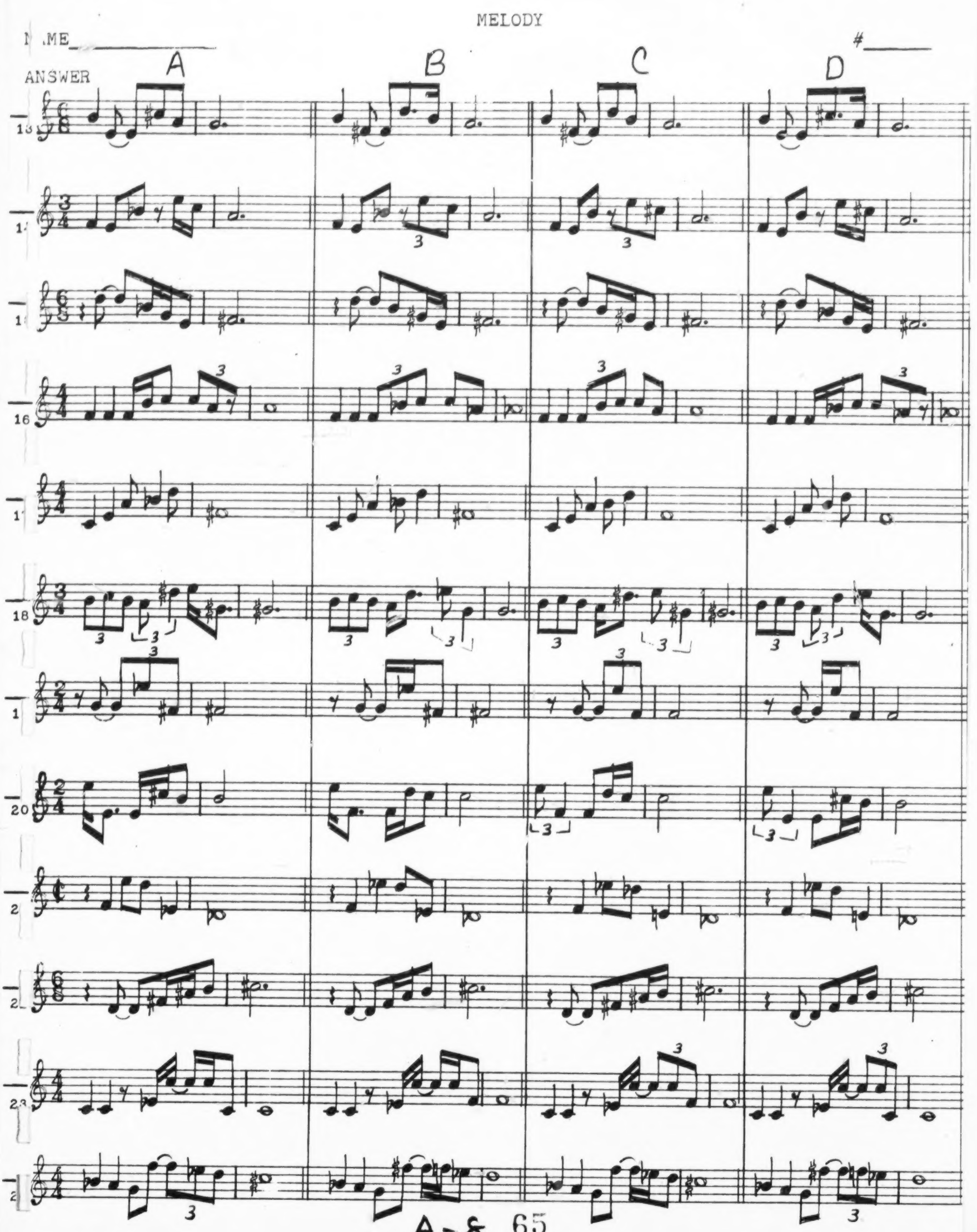
NAME_____ #_____

ANSWER

A B C D

MELODY

NAME_____ #_____

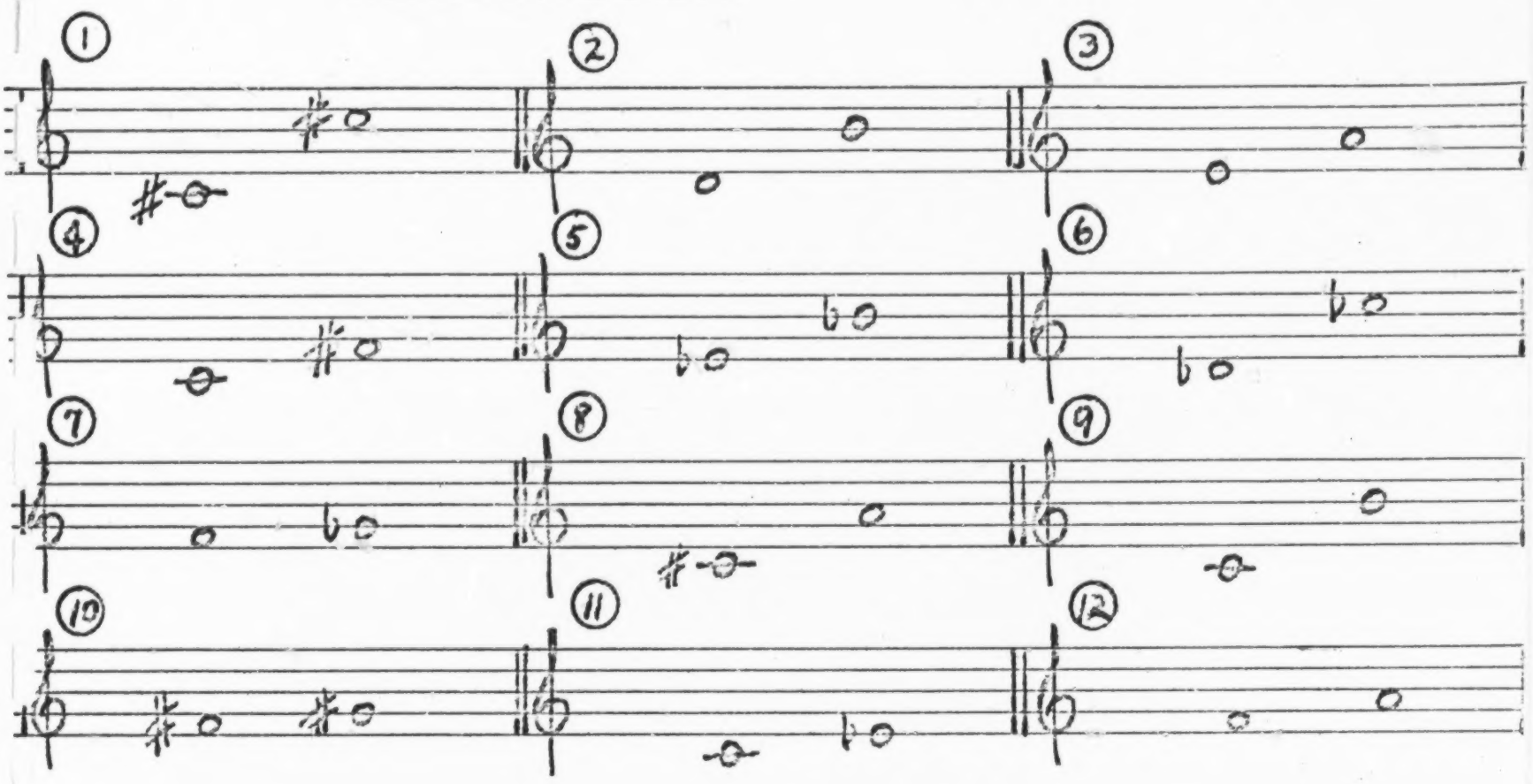
ANSWER

A B C D

A-8 65

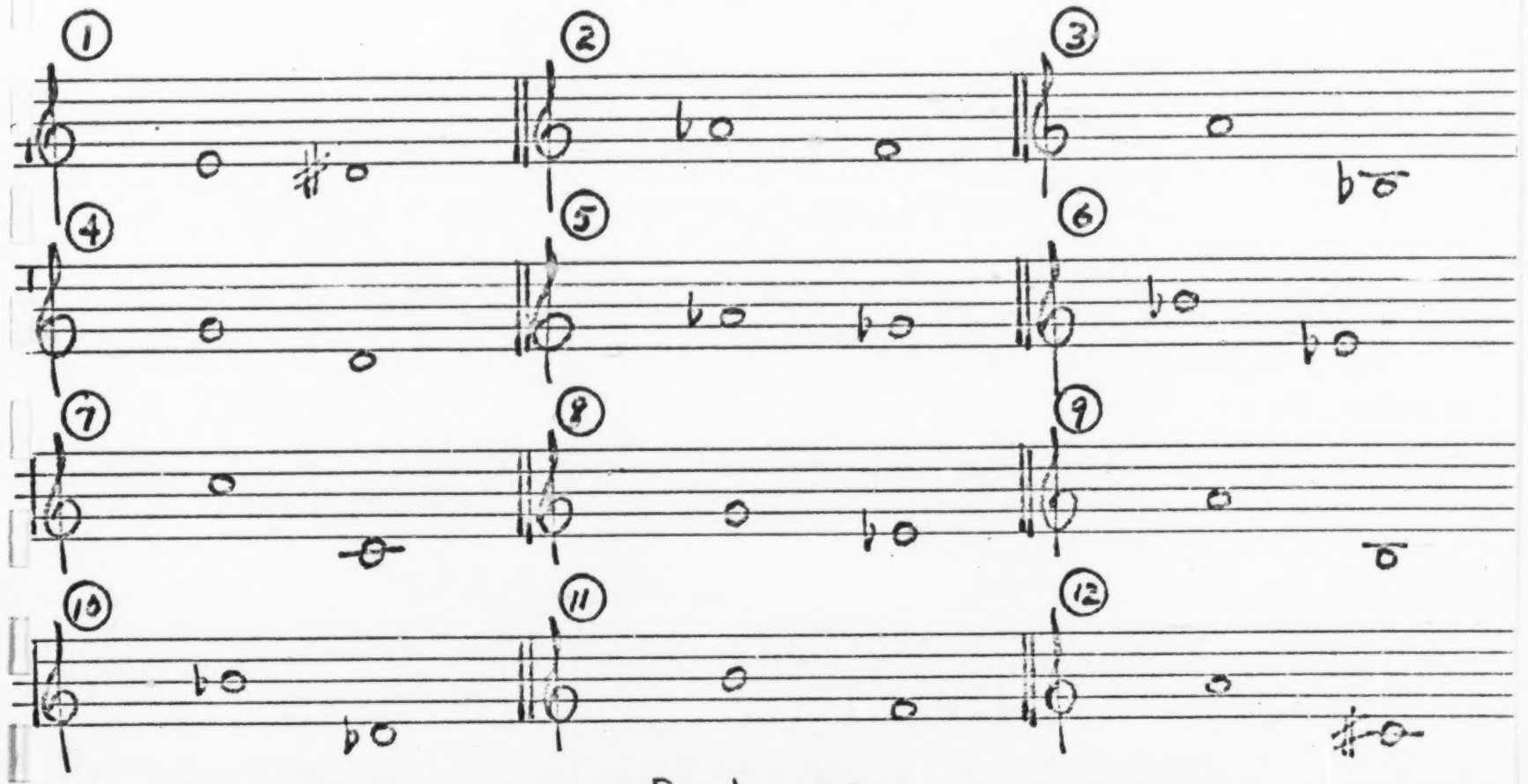
I. SIGHTSINGING MELODIC INTERVALS

Sing each interval using the neutral vowel "la". Rhythm is not a factor in this performance, that is, the two notes need not be of the same duration. However, each interval should be sung as a short melodic phrase without taking a breath between the two notes concerned.

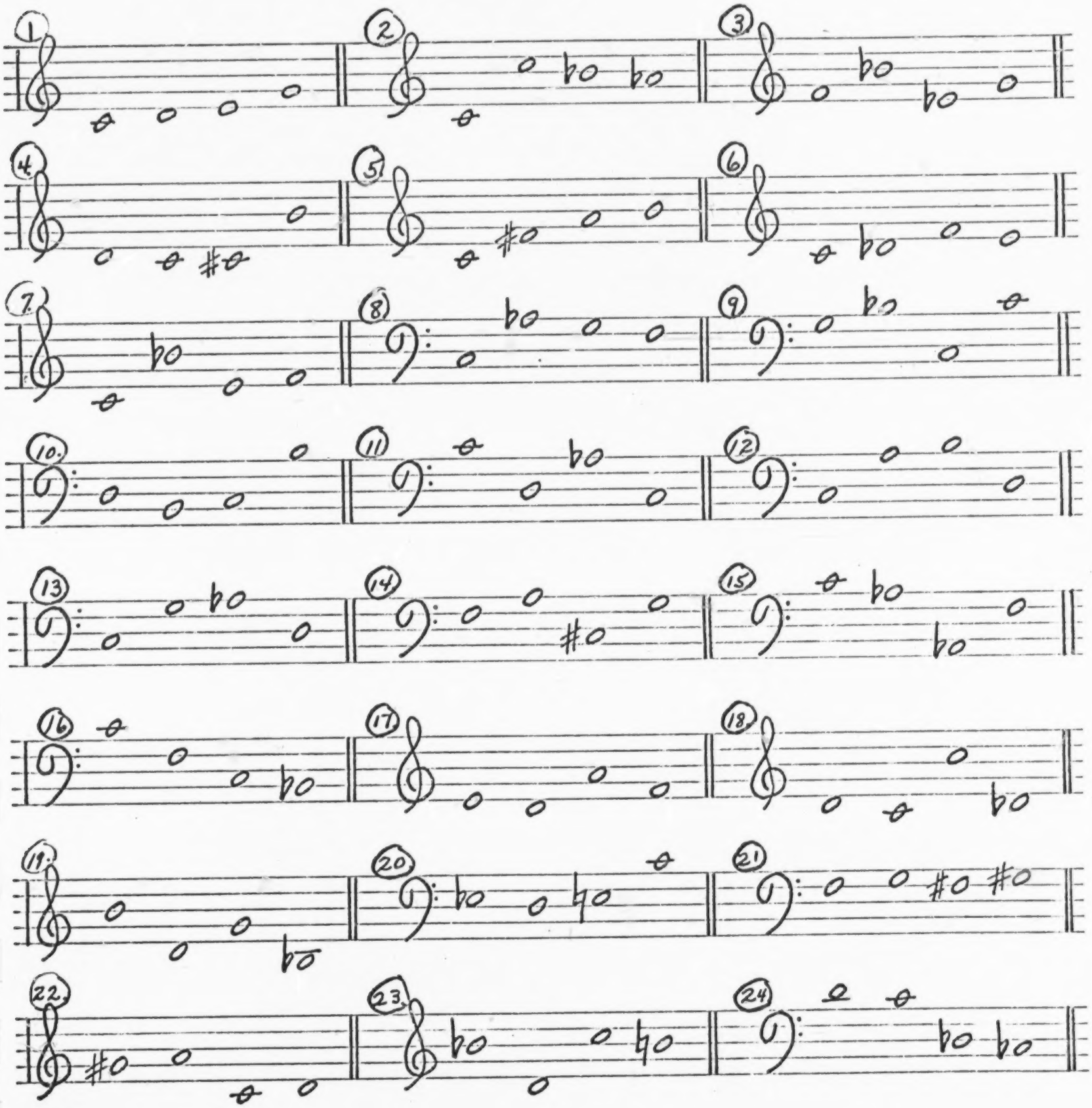
Upward melodic intervals



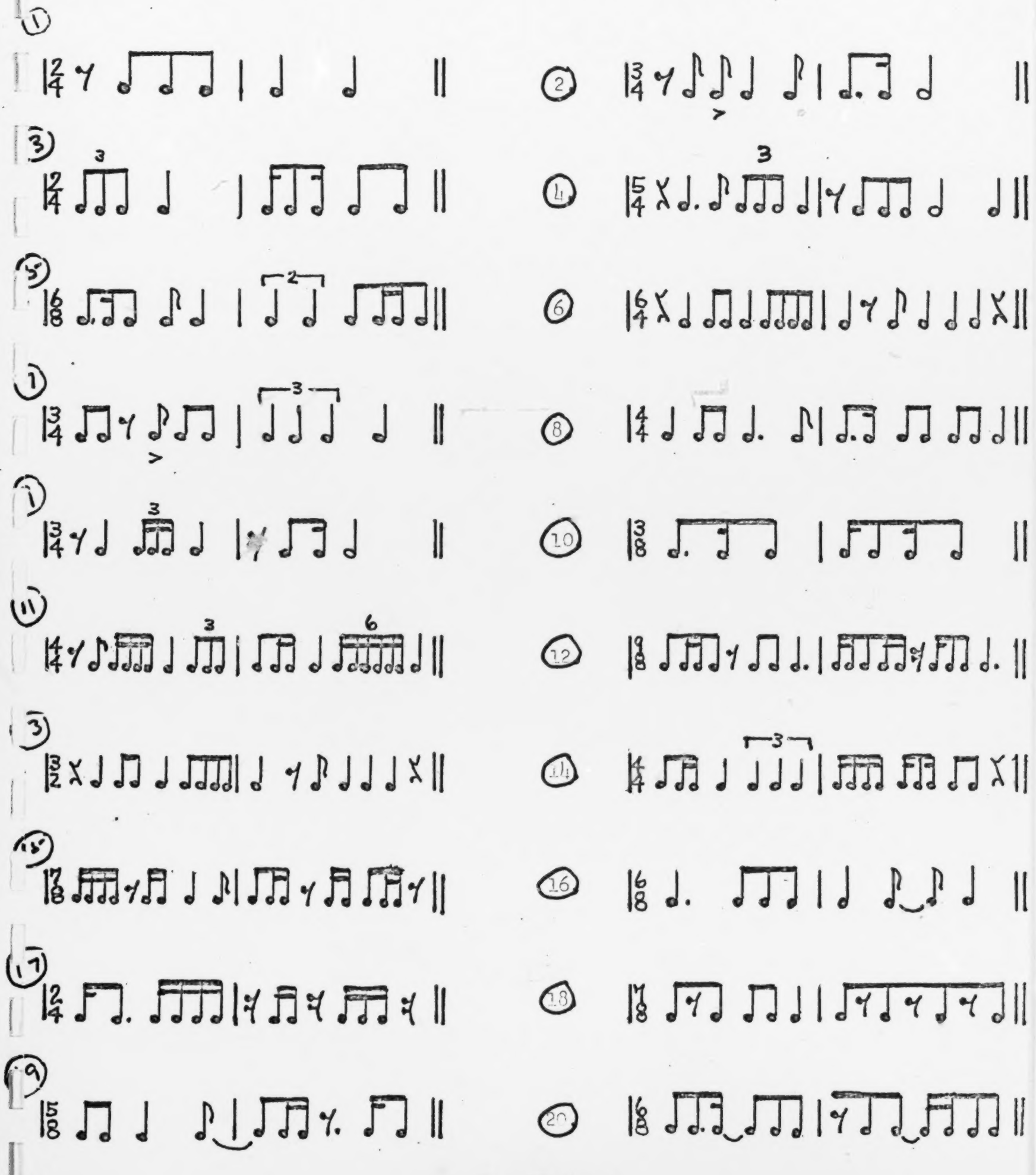
Downward melodic intervals

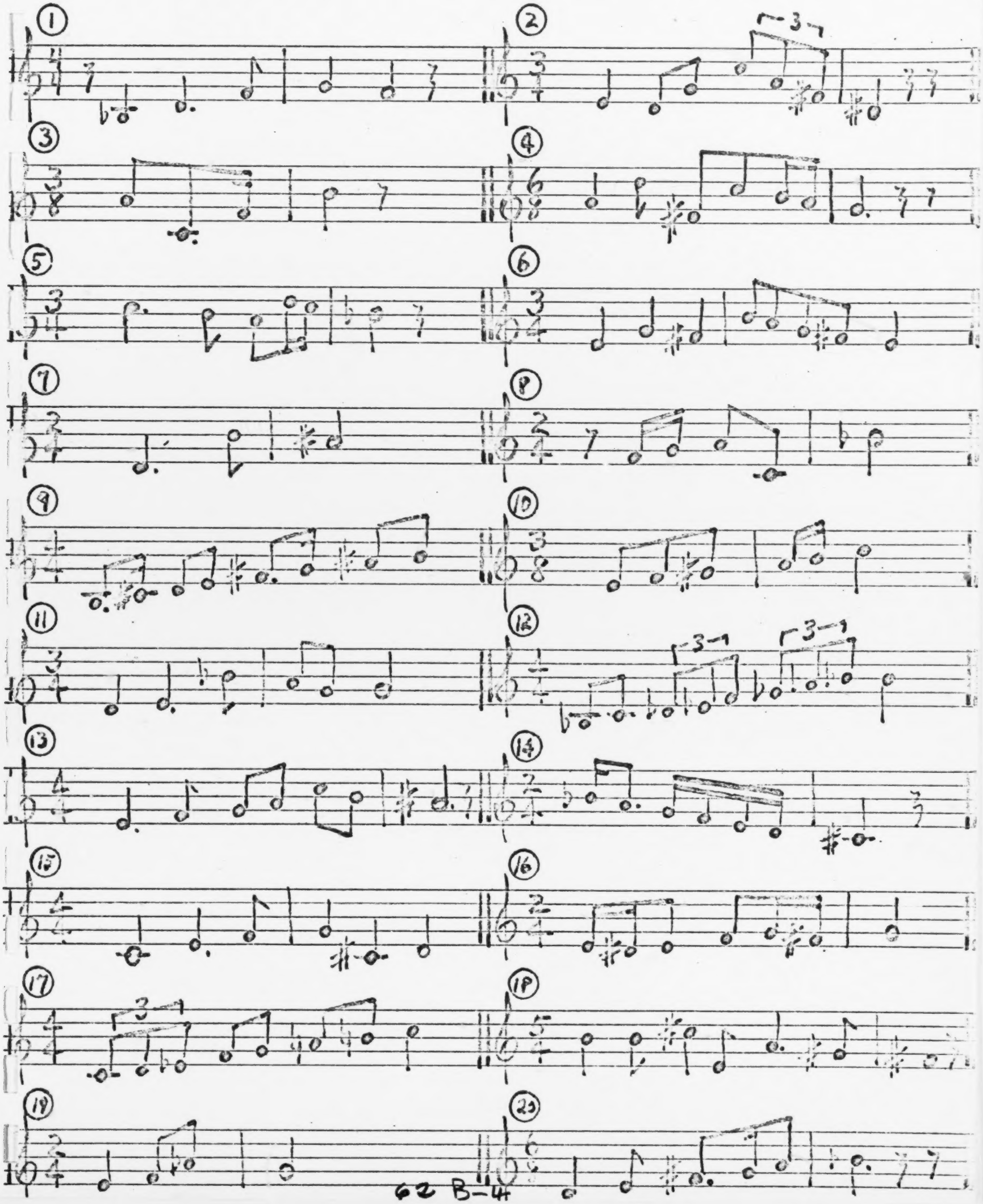


II. SIGHTSINGING PITCH PHRASES

Sing each four-note pitch phrase using the neutral vowel "la." Rhythm is not a factor in this performance, that is, the successive notes need not be sung using exactly the same duration. However, each four-note phrase should be sung in a single breath, that is, without stopping for a breath after once having started.

III. SIGHTSINGING RHYTHM PHRASES

Sing each rhythm phrase on either "la" or on "ta", as seems best in each instance. Pitch is not a factor in this performance.



IV. SIGHTSINGING MELODIC PHRASES

Sing each melodic phrase on the neutral vowel "la." Pay careful attention to the meter signature, to the rhythm values, and to the pitch notation used.